Author:
Ward, Daniel

Title:
Multiscale Modelling of Stem Cell Population Dynamics

Author:
Ward, Daniel

Title:
Multiscale Modelling of Stem Cell Population Dynamics

Multiscale Modelling of Stem Cell Population Dynamics

By

DANIEL TIMOTHY WARD

Department of Engineering Mathematics
UNIVERSITY OF BRISTOL

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of DOCTOR OF PHILOSOPHY in the Faculty of Engineering.

SEPTEMBER 2018

Word count: Forty five thousand five hundred

Abstract

Stem cells, with their propensity for pluripotency and self-regulation, have been the focus of intense research over the past decades. Understanding how the pluripotent phenotype is governed, across scales ranging from the cells' underlying genetic components to population-level interactions, is key to progressing from cultures of thousands of stem cells to maximal cell cultures, or entire tissues, from a small starting population. Stem cells are the precursors to all living cell types, existing during the embryonic development stage, but they can also be found throughout the bodies of all developed living creatures, for example within the bone marrow or residing at the base of an intestinal crypt. The differential behaviours of embryonic and adult stem cells, combined with complex underlying genetic regulation and environment specific dynamics, present a number of open and pressing questions. In this thesis, we present mathematical and computational modelling approaches which allow us to elucidate the key dynamics of three different stem cell populations, to shine a light on the differential effects of culture conditions on each cell type, as well as to highlight the usefulness of modelling for driving experimental research and development of culture protocols. In the first part of the thesis, we develop a delay differential equation model to capture the culture-dependent dynamics of a homogeneous population of human haematopoietic stem cells, and show that a key culture protocol component, epo, has two response phases in which it differentially affects cell proliferation and differentiation. In the second part, we introduce agent-based modelling as a method for capturing the population dynamics of mouse embryonic stem cells (mESCs) in different culture conditions. We showed that linking the MycN component of the mESC gene regulatory network to the cell cycle captures both the subcellular distributions of key proteins and growth dynamics *in vitro*. Finally, we develop a multiscale agent-based model of intestinal crypts, that couples ordinary differential equation modelling of subcellular kinetics to a cell-based description of cell movement, proliferation, and contact inhibition (CI). This enables us to recapitulate tissue level dynamics of intestinal crypts, as well as to present an alternative approach to describing the formation of the Wnt expression gradient. We showed that cross-talk between the Hippo and Wnt signalling pathways is able to affect CI and that CI is likely significantly reduced in intestinal crypt mutations. Together, this research shows the effectiveness of modelling, across physical and temporal scales, to recapitulate *in vitro* and *in vivo* stem cell dynamics, as well as to capture the contributions of key behaviours such as proliferation and differentiation to healthy and dysplastic population growth.

DEDICATION AND ACKNOWLEDGEMENTS

This thesis is dedicated to my nieces and nephew, Harry, Scarlett, Kyra and Freya.
I hope I can inspire you as you all inspire me.

I would like to thank, unequivocally and boundlessly, my supervisors Dr Lucia Marucci and Dr Martin Homer. For all the time spent reading through my work and providing invaluable feedback. Your tireless guidance and support has allowed me to grow both in my research and within myself.

I would also like to thank my collaborators, Dr Alexandra Gampel and Dr Alexander G Fletcher, for their knowledgeable and indispensable contribution and assistance. For the long hours spent conducting experiments, collecting and sharing their data, and for the hard work put in on our collaborative works, I would like to thank Ms Deborah Carter, Dr Elisa Pedone and Mr Simon Godwin.

Finally, I would like to thank my family. My parents, Karen and Laurie, for all the sacrifices you have made, the faith you put in me, and for the unwavering unconditional love and support, thank you I wouldn't be here without you. To my siblings, Simon, Emily and Hannah, so nobody feels left out here's your mention, I always know I can count on you, thank you all.

191064190768080688040792021093290198

AUTHOR'S DECLARATION

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: ..... DATE: ..... 20/03/2019 .....

TABLE OF CONTENTS

List of Tables

List of Figures

CHAPTER 1

BACKGROUND AND THESIS MOTIVATION

1.1 Introduction

Stem cells are cells that are able to differentiate into different cell types as well as to divide to produce more undifferentiated (stem)cells, in principle indefinitely. There are qualities that are common to all stem cells, in that they are unspecialised and are capable of self-renewal, maintaining a stem cell population. Stem cell function changes depending on the biological requirements of the system in which the stem cells reside, for example in the base of an intestinal crypt [63, 111], within bone marrow [143], or as part of the development stage of an embryo [90]. We distinguish between embryonic and adult stem cells; the former are derived at the early stage in the blastocyst and are pluripotent, meaning that they can develop into any cell type [18]. On the other hand, adult stem cells are limited to differentiating into cell types of their tissue of origin [184]. In the intestines, for example, adult stem cells actively renew the epithelial lining of the intestine at a rate of once every 4–7 days [6, 27], whereas the behaviour of stem cells found within bone marrow is dependent on key hormones [119]. The differences between adult stem cells residing in different locations throughout our body, and those cultured *in vitro*, suggest that we need to consider different methods for mathematical modelling, to allow for a better understanding of these seemingly similar but deceptively different cell populations.

The mathematical modelling of stem cell populations has evolved over the years, becoming more sophisticated and allowing those working in the area of mathematical biology to better replicate and thus gain understanding into stem cell population dynamics. Mathematical formalisms capturing stem cell dynamics can vary over physical scales, using ordinary differential equations (ODEs) to describe the temporal dynamics of protein concentrations within a cell and relevant gene regulatory networks (GRNs), or using partial differential equations (PDEs) to include spatial information to describe cell migration, proliferation (and division) and diffusion of

chemicals within a cell culture or tissue.

In this chapter, we discuss those biological factors that are common between various stem cell populations, as well as their key differences, and how we can tailor our mathematical models of these systems to best represent their dynamics *in vitro* and *in vivo* and, hopefully, better understand the factors that regulate stemness across biological systems. The stem cell types that we consider in this thesis are:

1. Haematopoietic Stem Cells;
2. Embryonic Stem Cells;
3. Intestinal Stem Cells.

We consider these particular stem cell populations as they allow for increasing levels of modelling complexity. Considering haematopoietic stem cells, cultured *in vitro*, and their proliferation and differentiation dynamics in response to growth factors provides a suitable system to consider a population level model. Embryonic stem cells can similarly be considered homogeneous at a population level, but have been shown to present stochasticity in their subcellular gene-expression level, in *in vitro* cultures, which suggests that an agent-based modelling approach can be combined with a stochastic element to elucidate single cell dynamics shrouded in the population-level only data and modelling approaches. Finally, looking at intestinal stem cells dynamics in tissues, we consider a system with combined spatio-temporal dynamics, with the physical structure of intestinal crypts combined with the subcellular dynamics governing the tissue behaviour. By considering each population of stem cells in turn we show the benefits of increasing complexity in mathematical and computational modelling of stem cell populations.

1.1.1 Human Haematopoietic Stem Cells

Human haematopoietic stem cells are the first cell-type belonging to the haematopoietic lineage, a class of cells which gives rise to all components of blood, whether found in the bone marrow, as part of blood (including platelets and red blood cells) or within tissue (such as lymphocytes or plasma) [143]. They are adult stem cells, residing in the human body within the medulla of the bone (the bone marrow), and are uniquely capable of giving rise to each of the different mature blood cell types and tissues. The process by which the body forms the blood cellular components is called haematopoiesis [132]. As haematopoietic stem cells are self-renewing they undergo asymmetric division: when they divide they produce a daughter cell which is also a haematopoietic stem cell, as well as a semi-differentiated cell along the haematopoietic lineage tree; in this way, the pool of stem cells is not depleted [163]. The semi-differentiated components of blood are not similarly self-renewing and as such, they divide symmetrically.

All differentiated blood cells can be divided into three lineages:

1. Red blood cells: oxygen-carrying cells;
2. Lymphocytes: immune system response cells;
3. Myeloid cells: involved in innate immunity and blood clotting.

It has been suggested that there are key hormonal dependencies to the effectiveness of haematopoiesis, with the possibility to drive the differentiation of stem cells towards particular cell types along the haematopoietic lineage tree, both *in vivo* and *in vitro* [163]. Whilst the typical methodology for culturing these stem cells *in vitro* to form red blood cells has included the application of particular hormones to aid in culturing a maximal number of cells, this process is not fully understood and as such the use of these hormones can not be considered an absolute requirement. It has been shown that the earliest of erythroid progenitors are responsive to a range of cytokines, including stem cell factor (SCF), interleukin 11 (IL-11), interleukin 3 (IL-3), thrombopoietin (TPO) and macrophage colony-stimulating factor (M-CSF) [51]. SCF, in particular, binds to its receptor KIT, a kinase which signals through a number of pathways including the PLC-γ, the phosphoinositide 3 (PI-3) kinase and Src kinases [180]. Additionally, the SCF cytokine has been shown to act alongside erythropoietin (epo) to govern the red blood cell progenitors proliferation and survival [180]. We are particularly interested in the use of the cytokine epo in *in vitro* culture of haematopoietic stem cells, known to be required for later stage erythropoiesis, but widely used throughout the culture of haematopoietic stem cells [177].

In this thesis we used mathematical modelling, fitted to experimental data, to capture the effects of hormones on human haematopoietic stem cell proliferation and differentiation capabilities *in vitro*.

1.1.2 Mouse Embryonic Stem Cells

Embryonic stem cells are pluripotent cells with an almost unlimited developmental potential, capable of maintaining their ability to differentiate into any type of cells ranging from nerve to heart to muscle [90]. They are derived from embryos, during the developmental stage before the embryo would typically implant in the uterus [115, 168]. Following fertilization, a series of cleavage divisions occur, with each of the cells (called blastomeres) of the cleavage-stage embryos remaining undifferentiated. In mice, after roughly five days of development, the cluster of cells begins to differentiate, with the outer layer committing to becoming part of the placenta (the trophectoderm) and separating from the inner cell mass (ICM) [90]. The ICM cells maintain the potential to differentiate into any type of cell of the body, but following implantation in the uterus, they are quickly depleted following rapid differentiation into other cell types, with a reduced developmental potential. If the ICM is removed from the embryonic environment, however, and subsequently cultured under appropriate conditions, the ICM-derived cells are able to maintain their full developmental potential, whilst simultaneously continuing to proliferate and replicate indefinitely *in vitro* [18].

Whilst the derivation of embryonic stem cells has been possible for decades, with mouse embryonic stem cells first reportedly derived in 1981 [115], and human embryonic stem cells derived in 1998 [168], still the interplay between factors and chemicals needed to robustly maintain pluripotency, or drive specific fate specification, is not completely understood. For example, the transcription factor Oct4 has been studied as a key marker for pluripotency for embryonic stem cells, with its expression at a critical level reported to be required for these stem cells to remain undifferentiated [14]. It has, however, also been shown to be insufficient by itself to maintain these cells in an undifferentiated state. More recently, Nanog has been identified as a key transcription factor for maintaining the undifferentiated state of embryonic stem cells, with the expression of Nanog decreasing rapidly as the cells differentiate [134].

With open questions surrounding the genetic dependency of embryonic stem cells in maintaining a pluripotent state, there are a significant number of genes involved in this process with the potential for cross-talk and interdependency of signalling pathways. The application of mathematical and computational modelling provides a means to investigate these possible interdependencies, and effectively determine possible ways to control the differentiation and state of pluripotency of these cells *in vitro*.

1.1.3 Intestinal Stem Cells

Intestinal stem cells are slowly dividing pluripotent cells residing in the walls of the intestines (both small and large) [151]. Along the intestinal walls are short indentations (crypts) which contain the stem cells at the crypt base, lined by semi-differentiated cells, with fully differentiated cells occurring at the top of the crypt at the crypt-intestinal wall junction [34]. The intestines are a rapidly renewing system, refreshed by the proliferation of these stem cells and their semi-differentiated progeny, every 4–7 days [6]. Whilst the system is similar in the small and large intestines, there are also key differences. In the small intestine, crypts are longer and contain additional Paneth cells: differentiated cells that also reside at the base of the crypt, interspersing the stem cells [170]. Additionally, the small intestine contains villi, finger-like protrusions which act to increase the surface area of the intestines (as well as potentially acting to alter intestinal fluid dynamics). These villi exist at the top of the crypts, and as such are replenished by the proliferating cells within the crypts, with fully differentiated epithelial cells forming the surface of the villi. In the large intestine, in the absence of villi, the top of the crypt meets with the intestinal wall [34].

Whilst intestinal stem cells have been studied for the past few decades, there are still significant questions surrounding the cellular signals which drive the regulation of cells existing within the crypt. For example, it has been shown that intestinal stem cells require a sufficient Wnt signal to proliferate [91], but the source of Wnt within the crypt is uncertain. Similarly, whilst the focus on recent studies has been on how Wnt signalling governs the proliferative and differentiative dynamics of the crypt, there have been suggestions of additional pathways

effectively combining with the Wnt pathway to affect behaviour linking cell proliferation and differentiation. For example, pathways known to affect cell proliferation and differentiation within the crypt include the Notch signalling pathway [15, 174], which has been shown to govern intestinal cell lineage specification [64] whilst also interacting with the Wnt signalling pathway [99]. Additionally, the Hippo signalling pathway has, more recently, been shown to play a role in the proliferation and differentiation of intestinal cells [25]. It has also been shown to interact with the Wnt signalling pathway [92, 175] to contribute to the cellular process of contact inhibition (the cessation of the cell cycle due to extracellular influence) [126, 188].

In particular, we would like to use mathematical representations of subcellular dynamics of intestinal stem cells in order to investigate the roles of the different signalling pathways in maintaining crypt homeostasis as well as the effects of mutations on disrupting this process.

1.1.4 Systems Biology of Stem Cells

In general terms, the research in this thesis is grounded in the principles of systems biology: computational and mathematical modelling of biological systems, that focusses on the complex interactions within the system as integral to the behaviour of the system. Such complex interactions are, for example, necessary for a proper understanding and prediction of health and disease. In particular, systems biology of stem cells can be described as the study of the emergent properties of a population of stem cells, and their constituent parts, through the use of quantitative experimental methods linked to predictive mathematical and computational models. These emergent properties are a result of the interactions, across spatial and temporal scales, of the stem cells when considered as a population. By considering the stem cell as a system, we can meld descriptions of the biochemical processes occurring within the cell with the spatial and structural components of the cell or population and the domain in which they reside [77]. This makes the understanding of the systems biology of stem cells an inherently multiscale problem. For example, we might consider the regulation of subcellular processes such as the transcription of key genes on a different temporal scale to the physical interactions and movements of the cell, but these processes may, in turn, govern behaviours at a larger physical scale such as cell-cycle dynamics, which subsequently alters the physical composition of the population and feed back into their subcellular dynamics. Figure 1.1 shows how the systems-level dynamics of cells interact with the temporal and physical behaviours at a larger scale, and how these different scales can then, in turn, be modelled.

The interplay between the multiple scales involved in stem cell systems can lead to the emergence of systems-level properties such as robustness, modularity, and population heterogeneity. As described in Mast, Ratushny and Aitchison (2014) [117], the goal of systems biology is to achieve more than a description of the individual components, and component properties, of a cell or population. The goal is instead to achieve a greater understanding of how information is transferred and subsequently interpreted by the cell [117].

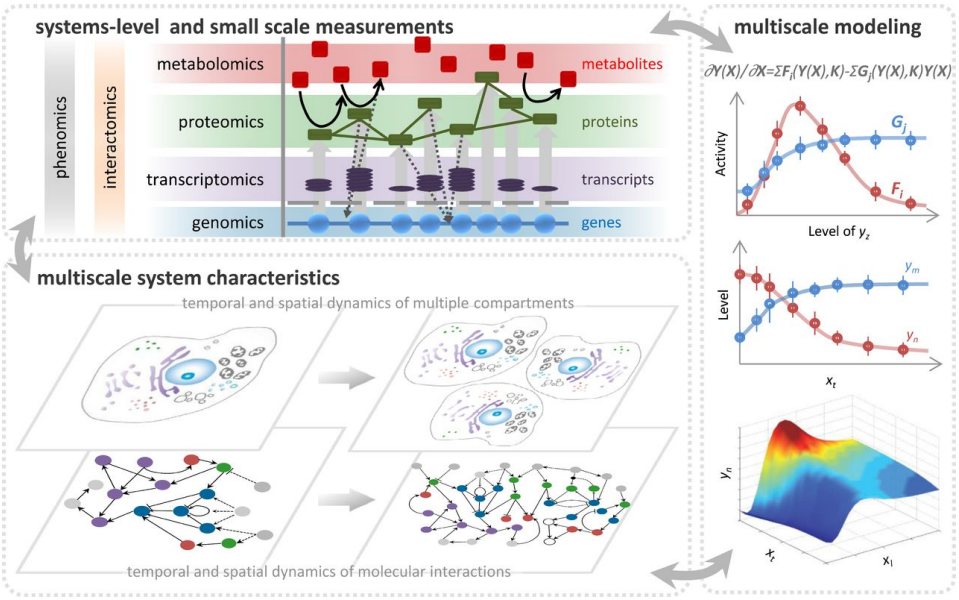


FIGURE 1.1. **Exploring the cell as a system.** Systems cell biology incorporates systems-level and small-scale measurements of information flow from all the components of the system and from all the hierarchies of the relevant temporal and spatial scale. Multiscale modelling approaches are used to infer missing parameters and visualize phenotypes as the function of multiscale system characteristics. The process is iterative, and the model is refined until it is predictive. Figure and caption reproduced from [117].

Typically, as presented in Figure 1.1, a systems biology approach includes a number of common elements [4]:

1. Exploratory data acquisition (through experimentation) and visualisation
2. Data integration and formulation of quantitative models
3. Testing of these models, alongside the hypotheses they generate
4. Further experimentation

These common elements can form an iterative process, by which experimental data feeds the development of a model, subsequently used to explain or describe a part of the system, which can be tested against further experimental data or used to suggest a change to the experimental method. The model can then be further refined in response to how well it captures features of the actual system alongside experimental data [4]. This process of developing models to better understand the complex interplay of cellular systems in governing population level dynamics has been effectively used over the years. For example, the eukaryotic cell cycle has been described using more than thirty models since 1970 [59], with the type of modelling applied ranging from

ordinary differential equation (ODE) representations [98], to delay differential equations (DDEs) [20], to stochastic models [158] as well as hybrids [9] of these approaches which can incorporate features of the ODE/DDE models along with stochastic elements.

In considering the scales of modelling that can be applied to systems biology of stem cells we can split these broadly into three categories: the microscale, the macroscale and the mesoscale. The microscale refers to subcellular behaviours, such as gene regulation and protein synthesis, whereas the macroscale refers to the interactions occurring at a cellular level such as cell-cell or cell-substrate interactions. Considering both of these scales at the same time results in the third scale, the mesoscale. Under this paradigm of systems biology, we have to consider which aspects of the stem cell type in question we want to capture, and in turn which scales we need to consider.

Stem cell systems biology provides us with a means to gain a quantitative understanding, through the development and application of mathematical and computational models, of stem cell heterogeneity. It makes it possible to quantify the response to drugs, pluripotency and differentiation potential, and is key to the development of robust *in vitro* culture protocols to allow us to understand and control stem cell dynamics in tissues and living systems.

In this thesis, we will consider the scales in turn, addressing problems of increasing complexity as described below.

1.2 Aims and Objectives

The aim of this thesis is to investigate the use of mathematical and computational modelling for furthering our understanding of adult and embryonic stem cell population dynamics. By increasing complexity, from a single cell level to population-level dynamics, across the systems considered we can determine the effectiveness of modelling in replicating the required cell behaviour, whilst showing how models can be used to test hypotheses and suggest new experimental or therapeutic approaches.

The main objectives of this thesis are

- Develop a mathematical model to describe individual stem cell response to external factors through population-level data of adult stem cell cultures;

- Use agent-based modelling to expand on single-cell modelling of embryonic stem cells cultured *in vitro*, to elucidate emergent dynamics in cell populations;

- Investigate an *in vivo* biological system at tissue level, using multiscale modelling to understand signalling pathways and cell cycle interplay, and the effect of genetic mutations.

1.3 Structure of Thesis

In Chapter 2, we begin by presenting an overview of methods which have been applied to model the different biological systems in each of the stem cell types we have considered in this thesis. In Chapter 3 we investigate the role of key hormones in the culturing of human haematopoietic stem cells, showing that the population can be effectively described by a delay differential equation model, and presenting results which indicate that the status-quo of haematopoietic culture may be incorrectly applying hormones during the early stages of differentiation. In Chapter 4 we investigate the dependency of mouse embryonic stem cell pluripotency on culture conditions, applying an agent-based model to couple gene regulatory network dynamics to the cell cycle and drug response; to recapitulate the different proliferative dynamics of cells cultured in different culture media. In Chapter 5 we investigate the intestinal crypt system, developing a novel cell signalling model to describe the cross-talk between two signalling pathways (Wnt and Hippo). The cell signalling model is then incorporated into a multiscale, multicellular framework, modelling the crypt as a whole, to further describe spatial dynamics in the crypt and formalise tissue behaviours in both healthy and dysplastic crypts. In Chapter 6 we conclude, by discussing the findings of our research and the suitability of the different mathematical and computational modelling approaches we used. We also discuss future avenues of research into stem cell population dynamics through the application of an ever-expanding repertoire of computational tools.

1.4 Publications

The work presented in Chapter 3 has been published as a research article: "Mathematical Modeling Reveals Differential Effects Of Erythropoietin On Proliferation And Lineage Commitment Of Human Hematopoietic Progenitors In Early Erythroid Culture" (Haematologica, 2016) [177]. The work presented in Chapter 4 has been published a research article: "An extended model for culture-dependent heterogenous gene expression and proliferation dynamics in mouse embryonic stem cells" (npj Systems Biology and Applications, 2017) [68]. The work presented in Chapter 5 has, as of writing this, been submitted as a research article: "Cross-talk between Hippo and Wnt signalling pathways in intestinal crypts: insights from an agent-based model" to PloS Computational Biology.

Chapter 2

Methods in Mathematical Modelling of Cell Population Dynamics

Stem cell populations exhibit complex behaviours, emerging across many scales physically and temporally. These behaviours include gene regulation at a single cell level, cell proliferation, and interplay of multiple gene networks. The application of mathematical modelling provides us with the means to better understand these complex systems, and the interactions across different scales, to elucidate the role of culture conditions on stemness in *in vitro* cultures, and to describe and predict stem cells dynamics in healthy and dysplastic tissues. In this chapter we discuss modelling methods specific to the three stem cell types considered in this thesis, providing the context for the application of specific methods used in each subsequent chapter. In particular, we consider the benefits and drawbacks of different methods, useful to make choices in our modelling work.

This chapter is arranged as follows. Section 2.1 details methods which have been used for modelling human haematopoietic stem cell (HHSCs) culture, with a focus on formalising key differentiation phases along the haematopoietic lineage, subsequently capturing the constituent contributions to population dynamics of cell proliferation and differentiation. Section 2.2 details the methods and tools involved in developing agent-based models of cellular populations, which we have applied to replicate *in vitro* stem cell cultures and *in vivo* tissue dynamics, focussing specifically on the agent-based modelling framework utilised in Chapters 4 and 5. Section 2.3 presents a range of modelling approaches that have been considered when attempting to replicate intestinal crypt systems, including implementations of agent-based models, as well as non-spatial

methods to capture healthy and dysplastic crypt dynamics. Additionally, in Section 2.3 we detail the specific modelling methods common to both agent-based models in Chapters 4 and 5, such as those used for cell-cycle modelling and replicating the mechanical behaviours of cells.

2.1 Modelling Cell Division Dynamics

In Chapter 3 we will consider the interplay between proliferation and differentiation of human haematopoietic stem cells (HHSCs), and develop a mathematical model to describe these competing biological processes and how they govern the population dynamics during the culturing process. It is possible to use mathematical and computational approaches to better understand these biological processes, to understand the complex interplay hidden within whole population data [40, 79, 121, 141]. In this Section we detail the methods considered to model the culturing process, to understand the benefits and drawbacks of each of the possible alternatives and the reasons for applying the methods used in Chapter 3. The approaches that we considered for modelling changing population dynamics through cellular division were:

1. Random birth-death model (RBD) [141];
2. Fixed-cycle model (FCM) [40];
3. Probabilistic birth-death (cyton) model (PCM) [79].

2.1.1 Random Birth-Death Model (RBD)

This is arguably the simplest of the possible modelling approaches that we can apply to describe the dynamics of a cell population, describing the process of division and death using exponential rates and ordinary differential equations (ODEs). Revy *et al* [141] proposed a model of T cell fate (a type of lymphocyte found along the haematopoietic stem cell lineage tree), that we summarise below. This model was developed to discriminate between the processes of cell proliferation and death in T cell populations cultured with added cytokines or with dendritic cells (antigen-presenting cells of the mammalian immune system). The model assumes that a cell is able to divide or die stochastically, and as such these processes can be described by division (α) and death (β) rates given a large enough, homogeneous, population. Based on these assumptions, the temporal dynamics of the growing cell population depends on both rates, whilst the relative cell counts of each generation depend specifically on the division rate. As such, the growth of a population can be described by the ODEs

$$\dot{N}_0(t) = -(\alpha+\beta)N_0(t), \tag{2.1}$$

$$\dot{N}_i(t) = -(\alpha+\beta)N_i(t) + 2\alpha N_{i-1}(t), \tag{2.2}$$

where $\dot{N}_0(t)$ is the number of cells in the first generation ($G(0)$) at time t, and $N_i(t)$ is the number of cells having completed i divisions within the population ($G(i)$). As previously described, the parameters $\alpha > 0$ and $\beta > 0$ are the rates for cell division and death, respectively. As the parameter values α and β are assumed constant over generations, the cell population is considered homogeneous; a simple extension to a heterogeneous population would be to use generation dependent rates for α and β. These equations can be solved, with the initial conditions $N_0(0) > 0$ and $N_i(0) = 0$ for $i = 1, ..., \infty$, to describe the changing population, for the first generation of cells as

$$N_0(t) = N_0(0)e^{-(\alpha+\beta)t}, \tag{2.3}$$

and for subsequent generations as

$$N_i(t) = \frac{(2\alpha t)^i}{i!} N_0(t), \qquad i = 1, ..., \infty. \tag{2.4}$$

This system remains, in practice, low dimensional as $N_i(t) \to 0$ when $i \to \infty$, as such earlier generations tend to zero as later generations arrive. The number of cells within the population changes exponentially, with the net growth rate $\alpha - \beta$, as the total cell population can be described by

$$N(t) = \sum_{0}^{\infty} N_i(t) = N_0(t)e^{2\alpha t} = N_0(0)e^{(\alpha-\beta)t} \tag{2.5}$$

Figure 2.1 shows the fitting results from Revy *et al.* [141], capturing the generational distribution (tracking number of cell cycles having occurred) at 24 hour intervals across the culture time course. The use of RBD modelling showed that the presence of dendritic cells during the culturing of T cells delayed T cell death, whilst simultaneously reducing proliferation (with the cell cycle of a T cell cultured with dendritic cells over 15 times longer than those cultured alone) [141]. This shows the capability of RBD models for elucidating differential values of cell proliferation and cell death from whole population count data. The RBD framework has also been used to investigate the extinction times of cancer stem cells compared to the normal haematopoietic stem cells [152], in order to propose an effective therapy to cause cancer stem cell extinction. They found that the resultant dynamics of the RBD model were heavily dependent on the difference between the birth and death rates, rather than on the individual rates. This makes the fitting of the rates less trivial as similar dynamics can occur at a range of rates, as long as the differences between them remain fixed [152].

Whilst the RBD is a simple but effective model for capturing population dynamics, there are several drawbacks. One downside to this method is that, as presented here, it assumes generational homogeneity. Additionally, it does not specify a minimum cell-cycle time for a cell

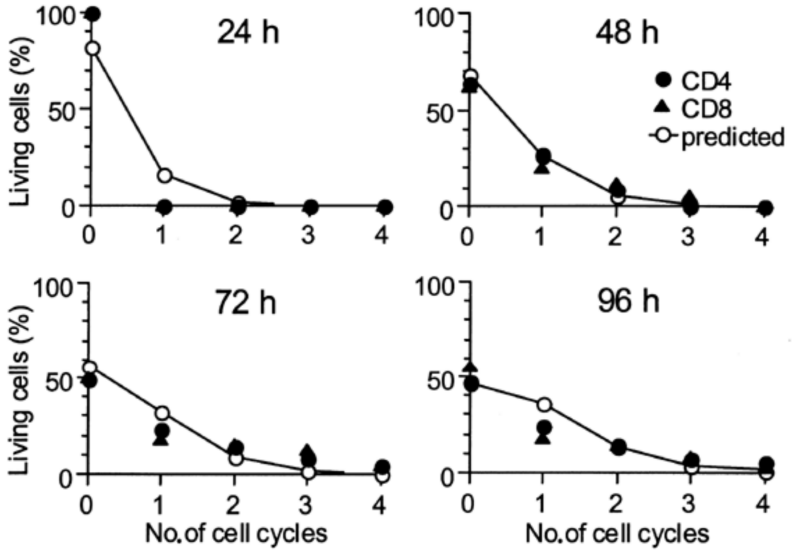


FIGURE 2.1. **Example of the application of RBD modelling to capture experimental culture dynamics.** $CD4^+$ or $CD8^+$ T cells, loaded with CFSE, were cultured with dendtritic cells for 1–4 days. RBD model, described in Equations 2.3-2.4, was fitted to population counts collected from CFSE data. Reproduced from [141].

instead relying on rates of proliferation. Finally, this model framework would not effectively capture the dynamics of a cell population in which there is a delay before the onset of proliferation, differentiation or cell death.

2.1.2 Fixed Cell-Cycle Model (FCM)

Building on the RBD model, an alternative modelling approach to encapsulate the exponential distribution of times to divide and die (as implied in the RBD model) is to assume a fixed cell-cycle time for cells within the population. The FCM model, as in Deenick *et al.*, [40] assumes that the time to first division (from $G(0)$ to $G(1)$) is sampled from a normally distributed random variable. Following this first generation, subsequent cell generations then divide following a fixed non-stochastic cell-cycle duration (Δ). As in the RBD model, cell death is considered to be independent of division, and as such is described by an exponential distribution with the parameter β_0 for undivided ($G(0)$) cells, and β for divided ($G(i), i > 0$) cells. The resultant model can be written as the following set of ODEs

$$\dot{N}_0(t) = -\frac{R(t)}{2} - \beta_0 N_0(t), \tag{2.6}$$

$$\dot{N}_1(t) = -R(t) - R(t-\Delta) - \beta_1 N_1(t), \tag{2.7}$$

$$N_i(t) = (2e^{-\beta\Delta})^{i-1} N_1(t-(i-1)\Delta). \tag{2.8}$$

where 2.6 describes the evolution of the first generation of cells, 2.7 describes the evolution of the second generation of cells and 2.8 describes the evolution of generations $i = 2,...,n$ where n is the number of cell divisions that have occurred. The function $R(t)$ is a recruitment function, describing the entry of cells into the first division stage [7]. Unlike in Deenick *et al.* where they propose a normal distribution for the first entry times, more recent experiments using tritiated thymidine uptake suggest that a lognormal distribution is a better fit [108]. As, during the S phase of dividing cells, thymidine is incorporated into the DNA, it is possible to determine the proportion of cells in the S phase at any given point during the experiment. This makes it possible to determine the time of entry into the first division for the cell population. With this in mind, the recruitment function can be modelled as

$$R(t;\mu,\sigma,\Delta_0,C) = \frac{C}{\sqrt{2\pi}(t-\Delta_0)} \exp\left(-\frac{(\log(t-\Delta_0) - log(\mu))^2}{2\sigma^2}\right). \tag{2.9}$$

Here, the parameters μ and σ describe the mean and the variance of the lognormal recruitment function and Δ_0 is an initial transient phase (such that the mean time for the first cell division is $\mu + \Delta_0$). The parameter C is set equal to $2\phi N_0$, where N_0 is the initial number of undivided cells in the population and ϕ is the precursor fraction; the latter indicates the number of cells which would have divided if death was discounted. The FCM described in Equations 2.6-2.8 can be generalized in order to account for a heterogeneous cell death [108]; in this case, for generations $i \geq 1$, death rate parameters can be defined as $\beta_i = \beta + m(i-1)$, resulting in the following

$$\dot{N}_i(t) = 2^{i-1} R(t-(i-1)\Delta) \exp\left(-\sum_{j=1}^{i-1} \beta_j \Delta\right) - 2^{i-1} R(t-i\Delta) \exp\left(-\sum_{j-1}^{i} \beta_j \Delta\right) - \beta_i N_i \tag{2.10}$$

for the ith generation of cells ($i \geq 2$). The equations for generations $i = 0$ and $i = 1$ remain the same (i.e. Equations 2.6, 2.7).

The FCM has been successfully used to investigate T cell regulation [36, 38, 40], by providing a quantitative framework to reconcile and explore experimentally questions of signal integration *in vitro*. It has been used to dissect the manner in which key cytokines (IL-2) promote T lymphocyte growth [40], elucidating the contributions of both differential cell death and cell proliferation in a population of T cells cultured *in vitro* with increasing concentrations of IL-2. Deenick *et al.* were able to show that the application if IL-2 during *in vitro* T cell culture does not alter the rate of

initial cell death in the population, but instead reduces the time taken for cells to first divide. Additionally, they were able to show that the application of IL-2 increases the rate of division of the cultured cell population. These results were not clear from considering the whole population data alone, and as such show the usefulness of the FCM in differentiating between the various cellular processes and their response to extracellular cues.

2.1.3 Probabilistic Birth-Death (Cyton) Model (PCM)

Unlike the assumptions of the previously described models (RBD & FCM), it has been suggested that the cellular control of cell division and death operate independently of each other, from observations obtained from several experiments [66, 79, 165]. As such, the time to divide and the time to cell death for each individual cell can be considered random variables. Hawkins *et al.* [79] named this approach the probabilistic cyton model (PCM). Cell division and death are considered independent of each other, and also independent between cell generations; it follows that the number of cells expected to have undergone a given number of divisions at any point in time can be determined from the probability distributions from which the process times were drawn.

This PCM model, as originally formulated by Hawkins *et al.* [79] describes the number of cells that have undergone i divisions that divide ($n_i^{div}(t)$) and die ($n_i^{die}(t)$) at a given time t. The first generation of cell dynamics are described by

$$n_0^{div}(t) = F_0 N_0 \left(1 - \int_0^t \psi_0(s)ds\right)\phi_0(t), \tag{2.11}$$

$$n_0^{die}(t) = N_0 \left(1 - F_0 \int_0^t \phi_0(s)ds\right)\psi_0(t), \tag{2.12}$$

where $\psi_0(t)$ and $\phi_0(t)$ are density functions which describe the probability that the generation 0 (i.e. undivided) cells divide and die, respectively, and F_0 refers to the fraction of cells which would be expected to progress to the following generation in the absence of cell death. Following from these equations, for an ith cell division, the functions $\psi_i(t)$ and $\phi_i(t)$ are again density functions which define the probability of division and death, respectively, and again F_i is the progression fraction in the absence of cell death, giving

$$n_i^{div}(t) = 2F_i \int_0^t n_{i-1}^{div}(s)\left(1 - \int_0^{t-s} \psi_i(s)ds\right)\phi_i(t-s)ds, \tag{2.13}$$

$$n_i^{die}(t) = 2\int_0^t n_{i-1}^{div}(s)\left(1 - F_i \int_0^{t-s} \phi_i(s)ds\right)\psi_i(t-s)ds. \tag{2.14}$$

From Equations 2.13 and 2.14 it is then possible to compute the cell numbers in each generation using

$$N_0(t) = N_0 - \int_0^t \left(n_0^{div}(s) - n_0^{die}(s) \right) ds, \tag{2.15}$$

$$N_i(t) = \int_0^t \left(2n_{i-1}^{div}(s) - n_i^{div}(s) - n_i^{die}(s) \right) ds. \tag{2.16}$$

From Equations 2.15 and 2.16 the progression of cells through cell division, with a possibility of dying, is described by the 'cytons' $\{\phi_i(t), \psi_i(t)\}$, which in turn specify the temporal probability of division and death for cells which have undergone i divisions, as well as the progression fractions F_i. It is assumed in [79] that the functions $\phi_i(t)$ and $\psi_i(t)$ are lognormal functions described by two parameters, the mean and variance. The simplest version of this cyton model assumes that $F_i = 1$ for all $i > 0$ and that death is negligible for divided cells (i.e cells in G(>0)) which is achieved by setting $\psi_i(t) = 0$.

The equations derived in Hawkins *et al.* [79] can be generalised. The number of cells that have undergone i divisions at a given time t is given by

$$N_i(t) = 2^i N_0 \left(\int_0^t r_i(t)dt - \int_0^t d_i(t)dt - \int_0^t r_{i+1}(t)dt \right), \tag{2.17}$$

where $r_i(t)$ is a density function which describes the probability of cells entering the i^{th} generation at time t. Similarly, $d_i(t)$ is a density function which describes the probability of cells in generation i dying at time t. Both of these density functions depend on the corresponding functions for the prior generations of the cells and, as such, they are typically defined via convolution integrals [8]. By choosing appropriate functions $r_i(t)$ and $d_i(t)$ it can be shown that the PCM is a generalization of the FCM described in Section 2.1.2 [8].

The PCM has been used to model the interplay of lymphocyte quiescence, proliferation and cell death, to answer the question of how lymphocytes (T and B cells) integrate receptor-mediated signals to affect these processes [79, 80]. It has been used to fit and analyse fluorescent division tracking data [79]; to further understand how receptor-mediated changes are able to modify the immune response, by incorporating the intrinsic variability in apparently identical cells as an essential feature of immune regulation [88].

Although the PCM accounts for a further level of stochasticity within the cell population, compared to the FCM we described previously, extending from a solely stochastic generation 0 to full stochasticity throughout the generations, it lacks the delay in the initial proliferation and differentiation of the cells within the population. This delay could be incorporated by having dynamically changing division and death probabilities. When extending this model to encompass a number of generations of cells across two subpopulations over the course of the simulation the application of this model would require the continuous generation of random variables to evaluate the changing cell numbers. Whilst this is not an inherently difficult problem to

overcome, due to the optimisation routines required to fit model parameters to experimental data the requirement for producing this significant set of random variables would reduce the efficiency of the optimisation and subsequent fitting of the model. If the drawbacks of this added computational load, and the risk of incorrect random variable seeding, outweigh the added effectiveness of fitting the experimental data, then we can not advocate for using this particular modelling approach.

2.2 An Agent-Based Modelling Framework (Chaste)

2.2.1 Agent-Based Frameworks for Systems Biology

Chaste (Cancer, heart and soft tissue environment) is an open source library coded in C++, for the simulation of physiological and biological mathematical models [122]. It was developed to provide open source methods for simulating complex models across scales in systems biology, applications include modelling cardiac electrophysiology [33, 136], modelling cell culture experiments [68], recapitulating epithelial layer dynamics [37, 47], and modelling intestinal crypt dynamics [48, 49, 60, 61, 105, 124]. The modular nature of the code library allows a large number of mathematical models to be formed more rapidly; Chaste includes solvers for ordinary/partial differential equations (ODEs/PDEs) as well as methods for linking these with bespoke input/output (IO) interfaces. Use of open source computational packages, such as Chaste, allows multiple users to check the base libraries of code, guaranteeing a heightened level of reliability and reducing the time spent having to troubleshoot or fix any broken or incompatible code.

Chaste is not the only open source modelling package for systems biology. Table 2.1 (reproduced from Gorochowski *et al.* [72]) presents a number of modelling tools available to develop agent-based models. It also describes the different capabilities of each of the various tools, including methods governing agent dynamics and features as well as environmental factors.

Whilst Chaste is not the only simulation package for systems biology, it does, however, combine many of the features of other, separate, simulation packages, allowing for agent-based models that also integrate continuum features such as ODE/PDE models of signalling pathways and chemical gradients. The industrial software engineering standards that were used in the design and implementation of the Chaste package contribute to both its reliability and also its plug-and-play nature. However, there are downsides in the use of Chaste, with typical use limited to Linux (and more recently MacOS) based machines. Still, the development of Chaste is such that new editions are continually being released, aiming to improve the multi-threaded performance of the code and allow running of larger multi-scale models on high-performance computing (HPC). It is the framework we shall use for the agent-based simulations described in this thesis, for reasons set out in the next section.

Name	Agent dynamics and features*										Environment				Language	Refs
	Simple rules	Advanced rules	ODEs	DDEs	Chemical equations	Stochastic dynamics	Motility	Chemotaxis	Cell replication	Cell morphology	2D	3D	Chemical diffusion	Complex objects†		
AgentCell					•	•	•	•				•			Java	[53]
BacSim‡		•			•	•			•		•		•		Obj-C/Java	[103]
BNSim	•	•			•	•	•	•	•			•	•		C++	[178]
BSim	•	•	•	•		•	•	•	•			•	•	•	Java	[72, 118]
CellModeller	•	•	•			•			•	•		•	•		Python$\oint$	[144]
Chaste	•	•	•	•	•	•			•	•	•	•	•	•	C++	[122]
CompuCell3D	•	•	•			•	•	•	•	•	•	•	•	•	C++/Python	[161]
DiSCUS	•	•	•						•	•	•				C	[86]
FLAME	•	•				•			•		•				Python	[69]
gro	•	•				•	•	•	•	•	•		•		C++	[96]
iDynoMiCS		•			•	•			•		•	•	•		C++	[106]
NetLogo	•						•		•		•	•			Scalla‖/Java	[154]
Organism	•	•	•						•	•		•	•		C++	[29]
RapidCell					•	•	•	•			•				Java	[176]
Repast HPC	•	•									•	•			C++	[32]
Repast Simphony	•	•									•	•			Java	[130]

TABLE 2.1. **Comparison of agent-based modelling tools.** Figure and caption reproduced from [72]. *Columns are defined as follows. Simple rules, a limited subset of commands are available to control agent behaviours; Advanced rules, access to a full programming language is provided to control agents; ODEs, agents can use ordinary differential equations to describe their internal state; DDEs, agents can use delay differential equations to describe their internal state; Chemical equations, cellular chemical reaction networks can be simulated; Stochastic dynamics, the internal state of an agent and the interactions with other agents can be stochastic, i.e. upon meeting another agent, there is a probability that they interact; Motility, agents can move freely within the environment and functionality to manage collisions/interactions is available; Chemotaxis, a realistic implementation of chemotaxis is available to control cellular movement; Cell replication, agents are able to replicate over time; Cell morphology, agents can take an arbitrary shape or have the option to take one of multiple predefined shapes.
† Ability to define solid structures within the environment that have arbitrary geometries.
‡ BacSim is no longer developed and has been superseded by iDynoMiCS.
$\oint$ Simulations are accelerated using the PyOpenCL library, which provides access to parallel computation on GPUs.
‖ Scala code is compiled into Java byte-code to enable full interoperability with Java tools and other JVM-based languages.

2.2.2 Framework Design

Chaste was designed in the C++, object-orientated programming language. A benefit being its suitability for applications which require efficient memory management and performance [122]. In addition, the use of this programming language simplifies the process of extending existing functionality, which may previously have been hard coded. Chaste was developed in a test-driven manner, where 'test code' is written before the 'source code', which then performs the function that is required. Once the test code is written, the source code is subsequently written which allows the test to pass. On completion of the modelling, the 'test' and 'source' code is committed to a central repository. It is then run alongside the main Chaste code to make sure the new code works properly, whilst also making sure that no functionality in the main trunk has been compromised. Additional tests are conducted nightly to check all standard tests for memory leaks, that the documentation for all the source code is complete, and every line of the source code is executed by at least one of the tests. Chaste is developed using an 'agile' development methodology, reducing the planning of code development in order to focus on working on limited coding goals in a reasonable time frame. This approach allows for faster development of working prototypes of the code instead of waiting until the entire code setup is complete. Significant time is then spent reworking existing code, including the class structures and interfaces, in order to improve the efficiency, readability and ease of re-use.

One of the key benefits of Chaste is that it provides libraries for code which is common to a large number of computational and mathematical biology problems. The core of the Chaste library can be split into the following sections

- **global**: this contains the code for basic mathematics (including the random number generation), time stepping, checkpointing (saving and loading simulations) using the Boost serialization library to allow for saving individual simulations, parallel vector classes, and code to handle warnings and errors;

- **io (input/output)**: this contains the code for reading, writing and conversion of various file formats, including modules to handle the HDF5 scientific file format, which enables distributed data to be collated and stored in a single file;

- **mesh**: this contains the code for linear or quadratic tetrahedral meshes; nodes, elements, boundary properties; mesh generation; mesh distribution using METIS/parMETIS; readers and writers for Triangle/TetGen, Meshalyzer, Cmgui and VTK (Paraview) formats;

- **linalg (linear algebra)**: this contains the code which uses Boost uBLAS and PETSc for vector and matrix operations;

- **ode**: this contains the code for defining ODEs; e.g. solvers, basic finite difference schemes, and the Sundials CVODE solver;

- **pde**: this contains the code for defining elliptic and parabolic second-order PDEs; parallel finite element solvers of generic coupled systems of PDEs (using mesh and linalg);

- **continuum mechanics**: this contains the code for solving compressible and incompressible general non-linear elasticity problems.

This core library contains all the computational components that are used by the two main application areas of the Chaste code: cell-based and heart modelling. Individual bolt-on projects are another of the benefits of using Chaste, as the majority of the models that have been published having used the Chaste simulation package are readily available to download.

The cell-based Chaste code provides a range of modelling frameworks that can be adapted for the individual needs of the coder and the biological problem. There are hard-coded functionalities for biological processes such as the cell cycle, cell death, intracellular signalling pathways and coupled PDEs. For cellular simulations, there are two main types of spatial model: on-lattice and off-lattice. We discuss the specifics of these model types in Section 2.3. Additionally, as previously mentioned, the Chaste simulation framework provides tools which allow for the investigation of cardiac behaviour. Heart models in Chaste are able to describe the electrical properties of the myocardium by biodomain equations, a set of coupled parabolic and elliptic PDEs describing the tissue as two separate, distinct continua, providing a description of the ionic transport across the cell membrane of the heart [33, 136].

2.3 Modelling Intestinal Crypt Dynamics

In Chapter 5 we detail an investigation into intestinal stem cells, and consider the role that they play in the wild-type maintenance of intestinal crypts, and in dysplastic events involved in the development of colorectal cancer. When attempting to model the dynamics of intestinal crypt turnover, the methods that we consider can be broadly split into two categories [39]:

1. **spatial models**: models that describe the spatial location of each cell; they can be further divided into on-lattice (i.e. grid) models and off-lattice (i.e. lattice-free) models

2. **compartmental models**: models which simply describe the transition between cell classes without taking into consideration cells relative positions in the crypt.

Models that fall into the spatial models category are typically implemented in an agent-based framework as they describe the positioning and movement of cells in a spatial manner, with the addition of cell-based methods to describe cell cycle behaviour and subcellular dynamics. We detail the various implementations of spatial models, used in this thesis, including the possible benefits and drawbacks.

2.3.1 Spatial Models

With the coordinated behaviour of cell populations playing a central role in the growth and renewal of tissues, the interaction between cells and their microenvironment governing behaviours such as movement and proliferation are significant features to include in models. By considering cell populations in a spatial manner it provided a means to compare *in silico* experiments with their *in vitro* counterparts. There are various possible implementations of spatial models, with each having benefits and drawbacks which require consideration before choosing which to employ. Within the broad domain of spatial models, it is possible to break these down into two further categories, on-lattice (ON-L) models and off-lattice (OFF-L) models. Figure 2.2 presents schematics for five different spatial models that we can consider, with the cellular automaton (CA) and cellular potts (CP) models belonging to the ON-L model category, while the overlapping spheres (OS), cell-centred Voronoi tessellation (VT) and vertex based (VM) models belong to the OFF-L category.

2.3.1.1 On-lattice Models

The first ON-L model we consider is arguably the simplest cell-based model, the Cellular Automata (CA) model, where each of the individual lattice sites can contain at most a single cell. Figure 2.2(a) shows schematically how this model is set up, with each individual cell represented by a different coloured lattice site. This system is typically evolved discretely with fixed time-stepping [109], with the new state/position of the cell determined by using either stochastic or deterministic rules as well as the state of the system at the previous time step. These rules can come in the form of preferential direction of movement, with cells occupying a lattice site able to travel to neighbouring cells represented by the arrows in Figure 2.2(a), competing with neighbouring cells which may also prefer to move to the same lattice site. A major benefit of this CA cell modelling approach is the computational simplicity, making it suitable for simulating large numbers of cells within a population. However, as a cell is only described by an individual lattice site its morphology is generic and biologically unrealistic.

Another type of ON-L model is the Cellular Potts (CP) model [73, 133, 162, 179], which represents each cell by several lattice sites instead of individual lattice sites. This representation is shown in Figure 2.2 (b), with the use of several lattice sites per individual cell in turn allowing for more realistic representations of cell morphology. It is a mesoscopic approach for simulating cell and tissue dynamics, combining larger scale effects seen in a macroscopic model whilst also tracking individual cell identities. It can be considered a framework for model building, as opposed to a specific biological model, as it is capable of representing the spatial interaction behaviours of several populations, from people to cells. One early use of the CP approach was to simulate the cell re-arrangement seen in cell sorting experiments, and investigate differential adhesion of cells driven by cell adhesion molecules [67]. A benefit of using a CP as opposed to a CA model is that it is possible to incorporate additional mechanical components such as cell-cell

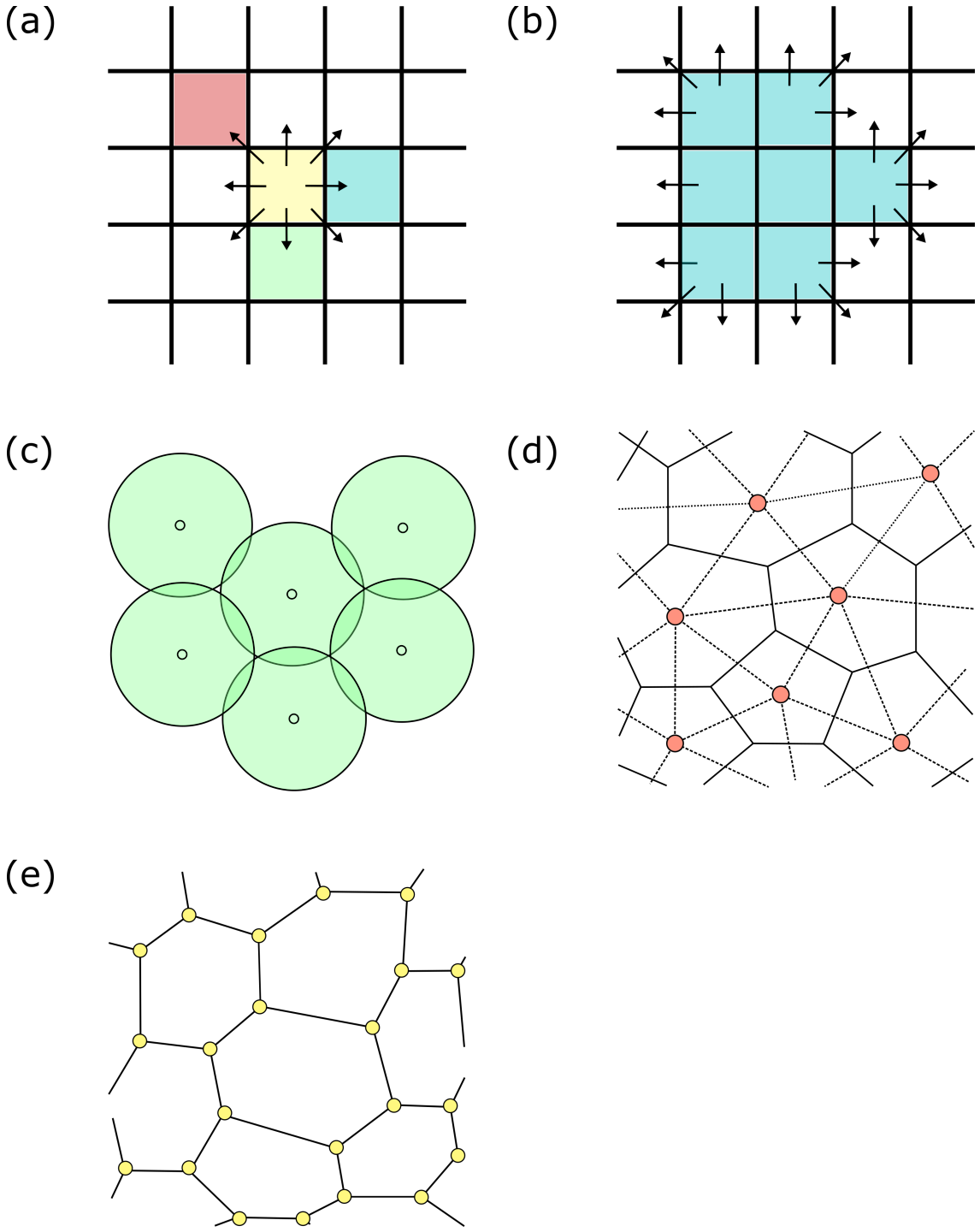


FIGURE 2.2. Schematics of cell-based spatial models. (a) Cellular automaton (CA). (b) Cellular Potts (CP). (c) Overlapping spheres (OS). (d) Cell-centred Voronoi tessellation (VT) model. (e) Vertex based model (VM).

and cell-substrate adhesion (due to possible differential cell-surface area between simulated cells composed of multiple lattice sites), as well as cell membrane tension and cell volume conservation.

The key input in the CP model is an *effective energy*, H, which forms the Hamiltonian of the system and governs the behaviour of the cells and their interactions with neighbouring cells. This Hamiltonian term subsequently contains the cell-cell and cell-*extracellular matrix* (ECM) interactions, as well as additional constraints that can determine individual cell behaviours.

Using a CP model, a collection of N cells can be described by first defining N spins (different states existing at the same energy level) $\sigma(i,j) = 1,...,N$, where (i,j) label a lattice site [162]. A cell σ consists of all sites in the lattice which then shares the same value of spin. These cells, therefore, form domains on a two-dimensional lattice. Each spin, and effectively each cell, interacts with its nearest neighbours on a square lattice, with the interaction energy of the corresponding interaction taking the form

$$H_{potts} = \sum_{(i,j),(i',j')neighbours} [1 - \delta_{\sigma(i,j),\sigma(i',j')}]. \tag{2.18}$$

where $\delta_{i,j}$ is the Kronecker delta. It is clear from Equation 2.18 that the two interacting spins corresponding to the two cells want to line up in order to minimize the interaction energy of the pair, with mismatched coupling between different cells having an energy of 1 and alike spins have an energy of 0. The simulated domain is evolved by following *Metropolis dynamics*: at each time step a new lattice site is chosen at random and its spin is changed from σ to σ' according to the Monte Carlo probability

$$P(\sigma(i,j) \rightarrow \sigma'(i,j)) = min\left\{1, exp\left(-\frac{\Delta H}{kT}\right)\right\} \tag{2.19}$$

for $kT > 0$, and

$$P(\sigma(i,j) \rightarrow \sigma'(i,j)) = \begin{cases} 0 & \Delta H > 0 \\ \frac{1}{2} & \Delta H = 0 \\ 1 & \Delta H < 0 \end{cases} \tag{2.20}$$

for $kT = 0$. Here, T is the absolute temperature, k is the Boltzmann constant, with the factor kT accounting for the amplitude of cell membrane (boundary) fluctuations, and ΔH is the energy gain produced by the changing of the spin. The H energy driving mechanism can subsequently be provided with additional components depending on the biological behaviour which has to be replicated.

Whilst the use of a fixed-lattice geometry description of the cell domain allows for increased efficiency when simulating large cell populations, especially with the CA model, there are benefits from shifting to an off-lattice approach. The latter modelling approach allows for a more detailed study of the morphology and subsequent mechanical effects of a population of cells.

2.3.1.2 Off-lattice Models

We consider three variants of OFF-L models, representing the cells in a continuous domain as opposed to a fixed-lattice domain. The two most common of the OFF-L model cell representations are the cell-centred Voronoi tessellation (VT) model [120, 147] and the 'overlapping spheres' (OS) model [45, 65].

The VT modelling approach labels each cell by the position in space of its nucleus (in the centre of the cell); this cell-centre is tracked over the course of a simulation. Each particular cell morphology is defined by the cell-centres of cells in the domain which fall closer to the centre of the cell in question than to the centres of any other cell, found by conducting a Voronoi tessellation of the cell-centres. A schematic representation of this approach is shown in Figure 2.2(d), with the cell centres represented by the red circles and the solid lines representing the Voronoi tessellation separating cells from their neighbours. Neighbouring cell centres are considered 'connected' to each other, with the connection typically represented by a spring force [120]; this method is discussed further in Section 2.3.1.3. The connections are determined by performing a Delaunay triangulation to connect those cell centres which share a common face from the Voronoi tessellation (dashed lines in Figure 2.2(d)).

Another common OFF-L model variant is the OS model, in which each cell is represented as a point centre with quasi-spherical morphology [45, 65]. In the OS OFF-L model, each spherical representation of the cell receives a force proportional to the volume of any overlapping regions between it and neighbouring cells or with a substrate surface. Figure 2.2(c) shows schematically how individual cells overlap with neighbouring cells, with the overlapping volume defining the forces applied to each cell. Whilst the VT model enables a more detailed representation of cell morphology, as well as providing a continuous description of the cell position within the domain, it is less suited for modelling cell-death. This is a result of representing the cells by a Voronoi tessellation: removing a cell from the system due to cell-death will result in the instant rearranging of the population such that every neighbouring cell becomes larger. Considering the OS representation, the cell death problem is overcome, as each cell morphology is not determined by its neighbours; thus upon the removal of a cell, neighbouring cells are able to freely move into the vacated region. However, there are also downsides to the OS approach. Cells are not typically spherical in reality, they generally have a more complex morphology which is excluded from an OS approach to modelling the cells. In addition, for each cell within the simulation which is overlapping, whether with a neighbouring cell or with a substrate surface, the volume of the overlap has to, in practice, be estimated. This estimation process is not perfectly accurate, resulting in possible discrepancies when applying forces or when implementing volume dependent behaviours. An additional benefit to the OS method, however, is its ability to shift from 2D to 3D, whereas using a VT model requires conducting Voronoi tessellations and Delaunay triangulations in 3D, which significantly increases computational costs. The OS representation has been extensively applied to researching tumour spheroid growth [65, 110], tumour monolayers

[45] and intestinal crypts [21, 48, 93, 94].

An alternative to the VT modelling approach which has been commonly used to describe epithelial sheets [62] is the vertex model (VM). Unlike the VT and OS models we do not consider each cell based on the position of its centre; instead, each cell is represented as a polygon, made up of individual vertices, which form the cell's membrane. This is shown schematically in Figure 2.2(e), with the yellow dots representing each cell membrane's vertices. In this framework, the cell vertices are moved according to a balancing of forces representing cell compressibility, cell-cell adhesion and cell-substrate interactions. Although applying a VM approach enables a more complex representation of the cell morphology, useful for modelling epithelial cells with complex geometries, for example, it suffers from similar drawbacks to the VT model. Again, it is not trivial to switch from a 2D to a 3D domain, although less complex than in the VT model as no tessellation is required. Additionally, the process of cell death requires the removal of vertices specific to multiple neighbouring cells, and so, in turn, additional rules to enable cells to shrink before being removed.

2.3.1.3 Cell-Centred Voronoi Tessellation (VT) Model

In considering the physical composition of our cell populations, in the agent-based modelling applied in Chapters 4 and 5, we apply the cell-centred Voronoi tessellation method proposed in Meineke *et al.* [120]. The model describes cells as a point centres (e.g. grey points Figure 2.3) corresponding to position vectors r. These points are attached to neighbouring cells by linear, over-damped springs. Assuming linear over-damped spring representations connecting the cell point centres, the movement of an example point centre (r_i) can be described according to

$$\frac{\alpha_i \eta}{\mu}\frac{d\mathbf{r}_i}{dt} = \sum_{j \in S_i}\left(||\mathbf{r}_j - \mathbf{r}_i|| - l_{ij}\frac{(\mathbf{r}_j - \mathbf{r}_i)}{||\mathbf{r}_j - \mathbf{r}_i||}\right) \tag{2.21}$$

where $\mathbf{r}_i$ specifies the position of the centre of cell i (e.g., blue shaded region Figure 2.3), $\mathbf{r}_j$ the cell centre of a neighbouring cell j (e.g., orange shaded region Figure 2.3), l_{ij} is the equilibrium length of the spring that connects cells i and j (equilibrium length is fixed throughout), S_i is the set of cells that lie adjacent to cell i (calculated from the Delaunay triangulation indicated in Figure 2.3), μ is the spring constant, and η the drag coefficient. The α_i term is a scaling factor which characterizes how the movement of cells differ; for example, that of mutant compared to normal cells. This means that $\alpha_i = 1$ for wild-type cells, and $\alpha \geq 1$ for mutants with increased "drag", which allows us to characterize how the movement of a mutant cell within the simulation differs from that of a healthy cell. As suggested from the experimental results by Sansom *et al.* (2004)[146], mutations in the Wnt pathway could potentially influence cell migration by increasing the cell-cell adhesion, due to an increase in cellular β-catenin which forms a bridge between adherens junctions and the cell cytoskeleton. This α scaling term is the focus of the mutant crypt experimentation in this model, providing a means to investigate how the possible advantage of a

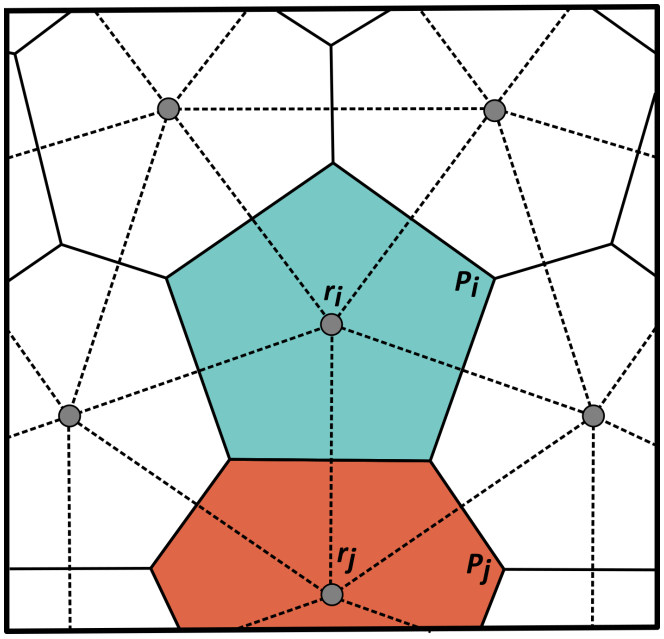


FIGURE 2.3. **Schematic diagram of the mechanical model applied in the agent-based modelling.** Grey points represent each cell centre. Pairs of neighbouring cells, identified by conducting a Delaunay triangulation, are considered connected by linear over-damped springs (represented by dashed lines). Solid lines represent the Voronoi tessellation, which divides the domain into individual cells (i.e regions shaded in blue and orange) such that all points P_i are closer to r_i than to any neighbouring centres.

mutant cell with increased cell-cell adhesion might increase its chances of colonising the crypt. In this model, it is assumed that all progenies of a given mutant cell share a common "drag" value $\alpha_i \equiv \alpha$, and so can be referred to in this case as simply the *adhesion scaling parameter*. The projected representation of our crypt therefore describes the position of our cells in two dimensions according to

$$(x, y) \in \mathbb{R}^2 : 0 \leq x \leq c, 0 \leq y \leq h, \tag{2.22}$$

where c and h denote the circumference and height of the crypt respectively, with periodicity imposed at the left- and right-hand boundaries (i.e. at $x = 0$ and $x = c$).

By modelling cells in this way we simplify the cells to a single point, with the morphology of the cell described by a Voronoi tessellation of the point centres. Although this method of describing the cell mechanics is relatively simple, compared to a vertex based model (modelling the boundary of individual cells as vertices, although still connected to neighbouring vertices in the same spring based manner), it has been used to accurately describe cell dynamics in a number of previous works [120, 124, 173]. When a cell divides following mitosis, a new cell-centre is introduced 0.1

cell diameters from the parent cell in a random direction. Following this introduction, the spring length between the mother and daughter cell is increased from 0.1 to 1 during the M phase of the cell cycle, in lieu of explicit cell growth.

2.3.2 Compartmental Models

Whilst the use of spatial models provides a key insight into the morphology of cell populations, as well as how individual cells interact within the global domain and with neighbouring cells, there are cases where there is no additional benefit to considering cells individually within a domain. The latter approach is preferable when representing large cell populations which can be considered homogeneous and spatial effects can be neglected.

Unlike the spatial models, compartmental models look specifically at the different cell types present in a domain of interest, and subsequently attempt to model the proliferation and differentiation of these cell types in a similar fashion to the models described in Section 2.1, neglecting the spatial placement of cells in the domain. Compartmental models allow us to describe gene-regulatory and signalling pathways to predict macro level cell counts. We consider two compartmental model representations of crypt dynamics below, presented in Boman *et al.* [16] and Garbo *et al.*[43] respectively.

The compartmental model described by Boman *et al.* [16] is formed such that there are predetermined subpopulations of epithelial cells along the crypt Y axis (height), with each of these subpopulations considered as a compartment. When a stem cell at the base of the crypt divides asymmetrically, it produces one each of generation g and generation $g+1$ cells. The generation $g+1$ is considered to be in a compartment that is further up the crypt Y-axis; in this way, the spatial organisation of the crypt is not modelled explicitly.

Figure 2.4 presents the crypt model design and modelling approach used to describe wild-type intestinal crypt dynamics. As a compartmental model, the crypt Y-axis is split into 82 increasing compartments, assuming that cells in later generations reside at increasing heights up the crypt. Splitting the crypt up in this fashion simplifies the model, as no explicit spatial constraints have to be considered. In addition, it reduces the description of the crypt to a set of equations which describe the differential transit through the cell cycle for the various cell types which reside within the crypt. One benefit of such a model is that it provides a focus on the cell cycle behaviour of the individual cell types within the crypt; stem cells at the base, partially differentiated cells towards the centre of the crypt and differentiated cells at the top of the crypt. The basic crypt set-up involves a set of equations which describe how cells within each crypt region pass through their cell cycle, subsequently creating cells in a later generation (assumed to reside further up the crypt). However, this formalism removes any capability to consider the spatial organisation of the crypt and does not allow for tailoring the behaviour of individual cells.

We next consider a more recent compartmental model, described in Garbo *et al.* (2010) [43], where a nonlinear mathematical model was applied to investigate the temporal evolution of cell

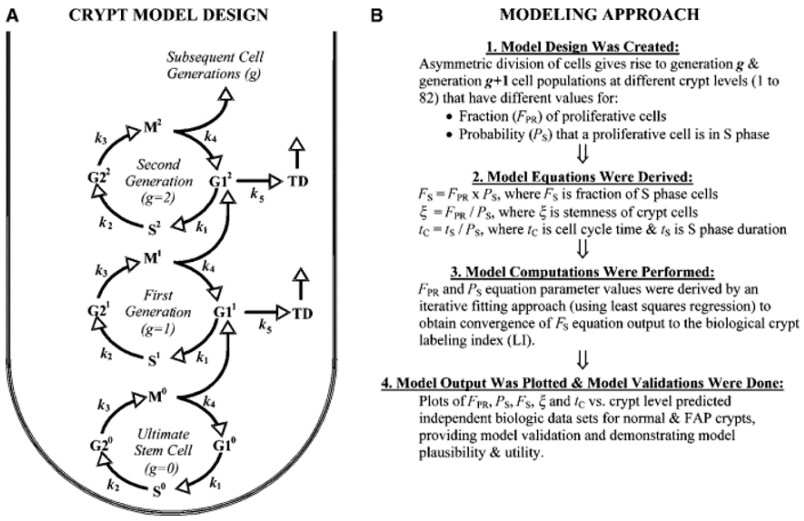


FIGURE 2.4. Crypt compartmental modelling approach described in Boman *et al.* (2001) [16]. (a) The crypt model is designed assuming cells in later generations reside at increasing heights up the crypt Y axis (height). This implicit spatial organisation reduces the problem to a generational representation of the crypt. Cell proliferation is modelled using a representation of the cell-cycle phases. (b) The modelling approach involved splitting the crypt Y axis into 82 individual levels with different values of proliferative cells and cell-cycle S phase residency. The model describes the crypt in terms of cell-cycle length, cell stemness and S phase residency. This model is then fitted to experimental data to match wild-type crypt dynamics. Adapted from [16].

populations in a colonic crypt. In this model, the crypt is split into three compartments containing, respectively, the base stem cells, the middle and semi-differentiated cells, and the upper and fully differentiated cells (Figure 2.5). As in the previous compartmental model, the spatial location of the cells is not considered.

In each compartment there are rates which govern the proliferative behaviour of the cells (α_3, β_3), the differentiative behaviour (α_2, β_2) and the apoptotic dynamics (α_1, β_1). In the final compartment, containing the differentiated cells, proliferation and differentiation do not occur and, as such, there is simply a rate of removal of cells, analogous to the sloughing of cells from the crypt into the intestinal lumen (γ). The temporal evolution of the population of cells within the crypt can be described in terms of the ordinary differential equations

$$\frac{dx}{dt} = (\bar{\alpha} - \alpha_2)x \tag{2.23}$$

$$\frac{dy}{dt} = (\bar{\beta} - \beta_2)y + \alpha_2 x \tag{2.24}$$

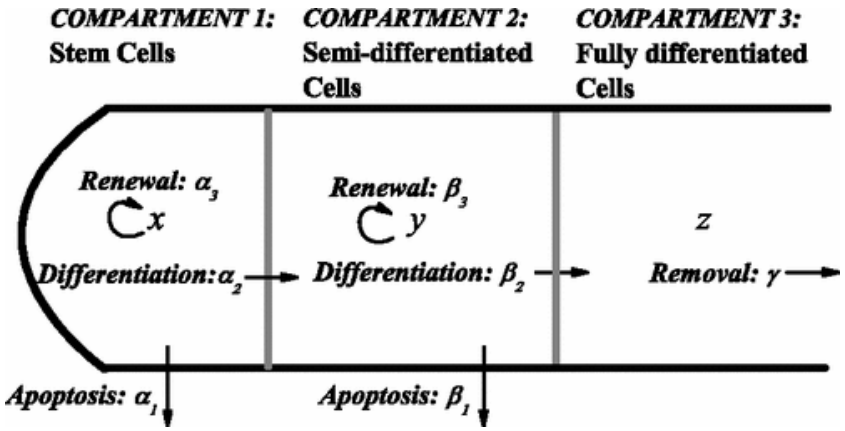


FIGURE 2.5. Schematic of the crypt model design applied in the compartmental crypt model described in Garbo *et al.* (2010)[43]. There are three different compartments, containing stem cells, semi-differentiated cells and fully differentiated cells. The arrows indicate the process which drive the cell behaviour, such as cell renewal, differentiation and apoptosis. Adapted from [43].

$$\frac{dz}{dt} = -\gamma z + \beta_2 y \tag{2.25}$$

where x is the number of cells in the stem cell compartment, y the number of cells in the semi-differentiated cell compartment and z the number of cells in the fully differentiated cell compartment, and $\bar{\alpha} = \alpha_3 - \alpha_1$, $\bar{\beta} = \beta_3 - \beta_1$. This system is described as being structurally unstable [43]; in order to obtain a more realistic description of the dynamics of the stem cells, the authors introduced the term $-kx^2 (k > 0)$ on the right-hand side of Equation 2.23, thus describing the time evolution of the population as logistic. In addition, the growth curves of all cell types were described by a logistic equation, with the rate of renewal of the population of semi-differentiated cells depending on the size of the population. As such, with these two additions, the rate of self-renewal of the cells in the second compartment y (β_3) is considered a sigmoidal function of the population size and as such can be described by

$$\beta_3(y) = \beta_3 + \frac{py^n}{q + y^n}, p > 0, q > 0, n \in \aleph \tag{2.26}$$

and so the temporal evolution of the crypt cell populations can be described by

$$\frac{dx}{dt} = (\bar{\alpha} - \alpha_2)x - kx^2, \tag{2.27}$$

$$\frac{dy}{dt} = (\bar{\beta} - \beta_2)y + \frac{py^{n+1}}{(q + y^n)} + \alpha_2 x, \tag{2.28}$$

$$\frac{dz}{dt} = -\gamma z + \beta_2 y. \tag{2.29}$$

Whilst the use of compartmental models is beneficial for obtaining insight into the processes driving population dynamics of the crypt on a mesoscopic level, the lack of spatial consideration in the model means that it is not suited for investigating the mechanisms which govern cell migration or additional physiological interactions between cells.

2.3.3 Cell-Cycle Modelling

As discussed in Section 2.2, constructing an agent-based model within the Chaste framework consists of defining three modules, the cell-cycle modelling module, the mechanical modelling module and the subcellular modelling module. In this section, we revise methods which can be used for cell-cycle modelling.

There are various possible models of the cell cycle that could be integrated into a multiscale model of a cell population. These range from the most simple, a fixed-cycle model which assumes a fixed time between subsequent cell divisions, to more complex representations describing the transition of cells through cell-cycle phases in terms of key gene dynamics.

In the fixed-cycle model, applied in a number of multiscale crypt models [21, 47, 60, 120, 124] and integrated into our crypt model in Section 5.3, it is assumed that there are two key states, as shown in Chapter 3, Figure 3.1. These states 'a' and 'b', representing the G_1, and S, G_2, M phases of the cell cycle respectively, are assumed to be stochastic and deterministic respectively. The result is that it is possible to allocate a specific set time to divide for each cell. To incorporate stochasticity into the population, preventing the synchronous division of cells, the 'a' state residency length for each cell can be drawn from a normal distribution $N(\mu, \sigma^2)$, where μ is the mean cell-cycle length and σ the cell-cycle standard deviation. The values of μ and σ can be found by fitting to experimental data [10, 27, 159]. By allocating fixed cell-cycle lengths from experimental data it makes the modelling simpler, requiring fewer parameters to be fitted and allowing the focus to shift to different aspects of the system, such as cell mechanics [48], or moving from a 2D to 3D domain [49]. Although modelling the cell cycle in this way reduces the complexity of the model, it does not allow for investigating the effects on proliferation of key genes that govern the cell cycle.

An alternative to the fixed-cycle model is to represent, in detail, key molecular components linked to cell-cycle progression; one such model was proposed by Swat *et al.* [160], which described the interactions between five molecular components, suggested by Kel *et al.* [100]. This is the cell-cycle description which is integrated into the multiscale model used in Section 5.4. Figure 2.6 shows the network of genes which control the transition of mammalian cells from the G_1 to the S phases of the cell cycle, and includes core proteins responsible for this G_1 to S phase transition: phosphorylated and unphosphorylated retinoblastoma proteins (pRB_p and pRB respectively),

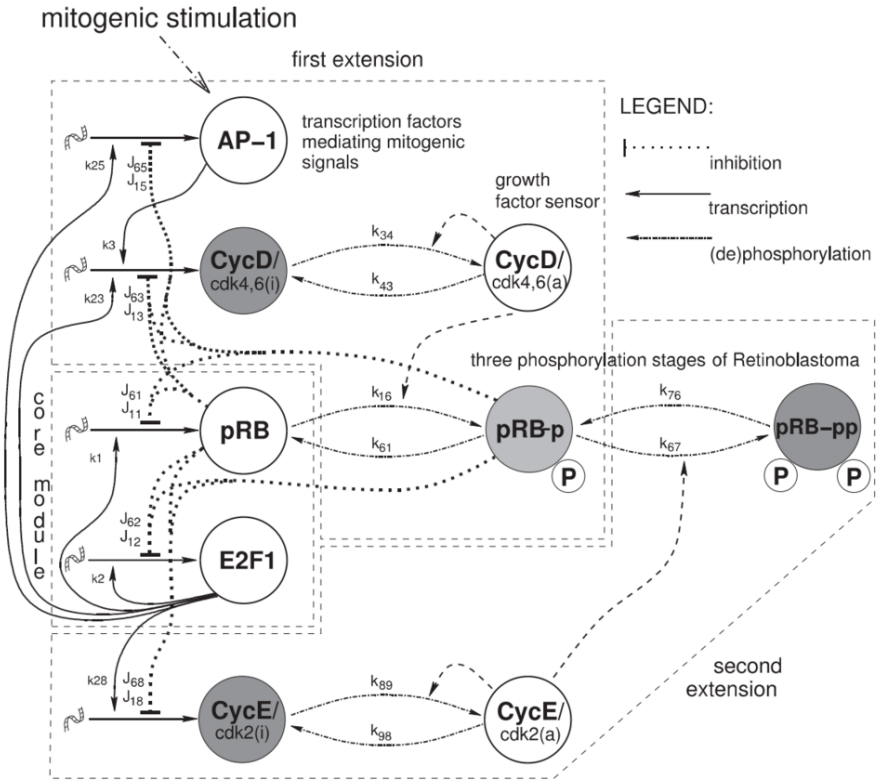


FIGURE 2.6. **Diagram representing a gene-network dependent model of cell-cycle G_1 to S phase transition.** Network incorporates key genes linked to mammalian cell-cycle progression, phosphorylated and unphospohrylated retinoblastoma (i.e pRB_p and pRB respectively), the transcription factor E2F1, active and inactive cyclin D kinase complexes (i.e $CycD_a$ and $CycD_i$ respectively) as well as the transcription factor AP-1. Diagram reproduced from Swat *et al.* [160]

TABLE 2.2. **Parameters for the Swat *et al.* [160] network description of the G_1 to S cell-cycle phase transition**. Fitting of parameters conducted and presented in [160], chosen to reflect experimentally known features of each molecular component.

$k_1 = 1$	$k_2 = 1.6$	$k_3 = 0.05$	$k_{16} = 0.4$	$k_{34} = 0.04$	$k_{43} = 0.01$
$k_{61} = 0.3$	$k_{67} = 0.7$	$k_{76} = 0.1$	$k_{23} = 0.3$	$k_{25} = 0.9$	$k_{28} = 0.06$
$k_{89} = 0.07$	$k_{98} = 0.01$	$a = 0.04$			
$J_{11} = 0.5$	$J_{12} = 5$	$J_{15} = 0.001$	$J_{18} = 0.6$	$J_{61} = 5$	$J_{62} = 8$
$J_{65} = 6$	$J_{68} = 7$	$J_{13} = 0.002$	$J_{63} = 2$		
$K_{m1} = 0.5$	$K_{m2} = 4$	$K_{m4} = 0.3$	$K_{m9} = 0.005$	$k_p = 0.05$	
$\phi_{pRB} = 0.005$	$\phi_{E2F1} = 0.1$	$\phi_{CycD_i} = 0.023$	$\phi_{CycD_a} = 0.03$		
$\phi_{AP-1} = 0.01$	$\phi_{pRB_p} = 0.06$				

the transcription factor E2F1, active and inactive cyclin D kinase complexes ($CycD_a$ and $CycD_i$ respectively), as well as the transcription factor AP-1 (Figure 2.6). The level of E2F1 has been shown to govern the G_1 to S phase transition [100].

The benefit of using this representation is that it details the interplay between crypt signalling and the cell cycle: crypt homeostasis is governed by Wnt signalling [107], which in turn is an inducer of CycD [78]. As the level of Wnt signalling is decreased, the amount of CycD expressed is reduced, and the expression of the downstream effector E2F1 is similarly reduced.

Based on diagram in Figure 2.6, Swat *et al.* [160] describe the network by six ODEs

$$\frac{d}{dt}[pRB] = k_1 \frac{[E2F1]}{K_{m1}+[E2F1]} \frac{J_{11}}{J_{11}+[pRB]} \frac{J_{61}}{J_{61}+[pRB_p]} - k_{26}[pRB][CycD_a] + k_{61}[pRB_p] - \phi_{pRB}[pRB], \tag{2.30}$$

$$\frac{d}{dt}[E2F1] = k_p + k_2 \frac{a^2+[E2F1]^2}{K_{m2}^2+[E2F1]^2} \frac{J_{12}}{J_{12}+[pRB]} \frac{J_{62}}{J_{62}+[pRB_p]} - \phi_{E2F1}[E2F1], \tag{2.31}$$

$$\frac{d}{dt}[CycD_i] = k_3[AP-1] + k_{23}[E2F1] \frac{J_13}{J_{13}+[pRB]} \frac{J_{63}}{J_{63}+[pRB_p]} + k_{43}[CycD_a] - k_{34}[CycD_i] \frac{[CycD_a]}{K_{m4}+[CycD_a]}, \tag{2.32}$$

$$\frac{d}{dt}[CycD_a] = k_{43}[CycD_i] \frac{[CycD_a]}{K_{m4}+[CycD_a]} - k_{43}[CycD_a] - \phi_{CycD_a}[CycD_a], \tag{2.33}$$

$$\frac{d}{dt}[AP-1] = F_m + k_{25}[E2F1] \frac{J_{15}}{J_{15}+[pRB]} \frac{J_{65}}{J_{65}+[pRB_p]} - \phi_{AP-1}[AP-1], \tag{2.34}$$

$$\frac{d}{dt}[pRB_p] = k_{16}[pRB][CycD_a] - k_{61}[pRB_p] - k_{67}[pRB_p][CycE_a] + k_{76}[pRB_pp] - \phi_{pRB_p}[pRB_p]. \tag{2.35}$$

In 2.30–2.35, the k's and ϕ's are non-negative rate constants, the K's and J's are Michaelis-Menten dissociation constants, and F_m is the mitotic stimulus [160]. In order to couple this cell-cycle model to Wnt signalling, Swat *et al.* [160] assumed that cyclin D is a Wnt target; as such Equation 2.32 can be rewritten as

$$\frac{d}{dt}[CycD_i] = \sigma k_3[C_T] + k_{23}[E2F1]\frac{J_13}{J_{13}+[pRB]}\frac{J_{63}}{J_{63}+[pRB_p]} + k_{43}[CycD_a] - k_{34}[CycD_i]\frac{[CycD_a]}{K_{m4}+[CycD_a]}. \tag{2.36}$$

In Equation 2.36 the value of $[C_T]$ represents the total level of β-catenin transcription factor complexes expressed in the subcellular Wnt model, and replaces the $AP-1$ term in Equation 2.32; as such we no longer need Equation 2.34.

The incorporation of the scaling factor, σ, is necessary to account for the discrepancy between the levels of β-catenin and AP-1 present in the cell (with Swat *et al* suggesting 8×10^{-3}nM $< [AP-1] < 1,24$nM, compared to Lee *et al.*[107] suggesting 25nM $< [\beta - catenin] <$ 155nM).

Table 2.2 presents the parameters used in [160], chosen to reflect experimentally known features of each molecular component of the network.

Chapter 3

Mathematical Modelling of Erythropoietin Effects on Human Haematopoietic Stem Cell Dynamics

We begin our research into stem cells by considering human haematopoietic stem cell (HHSC) dynamics. HHSCs belong to the haematopoietic lineage, a class of cells that make up the cellular components of blood. The differentiation lineage (i.e. irreversible differentiation of cells to a specific cell type) of HHSCs to fully-fledged, mature oxygen-carrying red blood cells is well understood, in the sense that we know the differential paths that cells can take when differentiating. The chemical requirements required for robust cell expansion protocols are still not completely understood. One of the key chemicals included in HHSCs culture protocols is erythropoietin (epo), a hormone which is essential for the production of mature erythroid cells promoting both proliferation and survival [102]. However, the role of epo in influencing lineage commitment of haematopoietic cells, both stem and progenitor, is unclear and of significant interest. A study has suggested that being exposed to high levels of epo can bias cell commitment towards the erythroid branch of the haematopoietic lineage tree over the myeloid lineage, both *in vivo* and *in vitro* [75].

In order to better understand HHSC epo-dependent lineage restriction, we developed a mathematical model that captures the specific proliferative and differentiative dynamics of the population, and fit the model to experimental data. The latter was gathered by collaborative experimentalists in the School of Biochemistry (University of Bristol), and involved culturing

HHSCs over several days with/without epo. Fitting our mathematical model to experimental data allowed us to better understand this key branch of the HHSC lineage tree and the role that the hormone epo plays, further addressing the question of whether the formation of lineage-restricted erythroid progenitors is coupled to cell division.

In Section 3.1 we provide the background biology, and general motivation for this research. In Section 3.2 we describe the experimental methodology and results gathered by our collaborators. In Section 3.3 we develop the mathematical model for the HHSC culture, to capture their proliferative and differentiative dynamics. In Section 3.4 we detail the model fitting results. In Section 3.5 we provide sensitivity analysis of the fitted model. In Section 3.6 we use our model to analyse the effects of epo on the culturing of HHSCs and to determine what effects it may have on the proliferation and differentiation of these cells. Finally, in Section 3.7 we discuss the results of our modelling of HHSCs, providing possible avenues for future research as well as indicating the drawbacks to our modelling approach.

The experimental data presented in Section 3.2 was collected by collaborators Dr Alexander Gampel and Ms Deborah Carter in the School of Biochemistry, Faculty of Medical and Veterinary Science, University of Bristol. The work detailed in this Chapter has been published as a research article: "Mathematical Modeling Reveals Differential Effects Of Erythropoietin On Proliferation And Lineage Commitment Of Human Haematopoietic Progenitors In Early Erythroid Culture" (Haematologica, 2016) [177].

3.1 Background

3.1.1 Motivation

Blood, in its many forms (including red cells, white cells, platelets and plasma) is in constant demand in the UK's NHS. It is required to keep millions of people alive, replacing blood lost due to injury or during surgery. In the UK, over 1.71 million donations are received per annum, whilst this is a significant amount it does not come close to the 2.1 million donations required by the NHS for all surgeries and life-saving treatments performed on the millions of patients each year [127]. On top of this, each donation of blood has a shelf life. Red blood cells, in particular, can only be stored for a period of 35, whilst white platelets can only be stored for 7 days [30]. This means that the NHS requires a consistent supply of blood in high quantities throughout the year. This gives rise to a significant challenge with two obvious solutions: to either convince increasingly large numbers of people to donate blood on a consistent basis, or to produce more blood in labs.

A major outstanding challenge for biochemists is to develop robust culture systems for mass amplification of red blood cells from haematopoietic stem and progenitor cells (HSPCs) for transfusion. Current culture systems are able to generate reticulocytes, the precursors to erythrocytes, which fully mature in the bloodstream. However, it is not currently feasible to amplify the starting HSPC population to produce even close to a unit of blood [55]. One possibility

would be to delay the maturation of the starting population whilst also maintaining high levels of proliferation. However, this is not as simple as it sounds, with proliferation and maturation coupled both temporally and mechanistically, such that the transcription factors that regulate the maturation and differentiation of the stem cells also control cell cycle regulation. If we can further understand the role of extrinsic factors on cell maturation and proliferation we can attempt to come up with ways to lower the costs involved and perhaps create a world in which blood donations no longer provide the majority of the blood required on a daily basis.

3.1.2 Cell-Cycle Biology

In order to better understand and subsequently model the cell population dynamics, through proliferation and maturation processes, we incorporate features of the cell cycle. The cell cycle refers to the process by which an individual cell grows, replicates its DNA, and then divides producing two daughter cells.

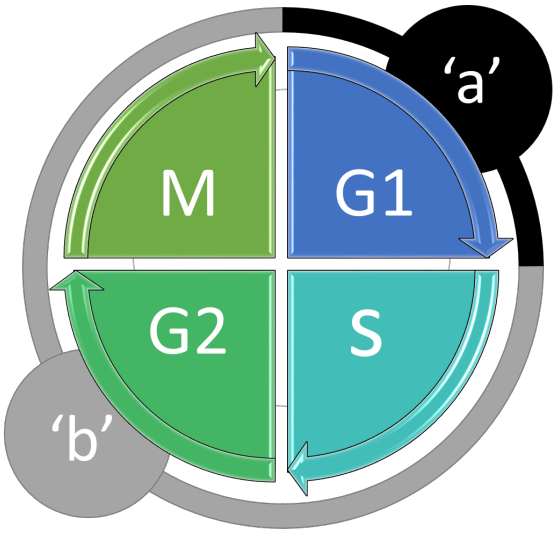


FIGURE 3.1. **Diagram illustrating the transitions of the stages that occur within the cell cycle.** These are Gap 1 (G1), Synthesis (S), Gap 2 (G2) and M (Mitosis). During G1 the cell increases in size before it enters S where DNA replication within the cell occurs. Once DNA replication has taken place it then enters another gap phase G2 where it continues to grow in size before finally entering M where the cell stops growing and divides into 2 daughter cells [28]. Also illustrated is how the stages of the cell cycle relate to the compartments of our adapted Smith-Martin model ('a' & 'b') [155], see Section 3.3.2.

The cell cycle is made up of 4 phases, Gap 1 (G1), Synthesis (S), Gap 2 (G2), and Mitosis (M) [28], depicted schematically in Figure 3.1. During the G1 phase, the cell grows in size, preparing for DNA synthesis in the later stages of the cycle. DNA synthesis then occurs during the S phase

of the cycle, followed by the G2 phase, another gap phase where the cell continues to grow in size. Once the cell has completed the G2 phase, cell growth ceases and the cellular energy is focused on the division of the cell into two daughter cells. Cell division then occurs following the completion of the M phase, during which the replicated chromosomes have been separated into two new nuclei.

Due to the complex nature of the cell cycle in mammalian cells, there are key checkpoints that have to be passed before the cell can continue through the cycle [160]. These checkpoints occur following the gap phases of the cell cycle (G1 and G2) and at the end of the mitosis (M) phase. The checkpoint following the G1 phase is the point at which the cell becomes committed to entering the cell cycle, which occurs when the cell activates cyclin-CDK-dependent transcription [28], which in turn promotes entry into the S phase of the cycle; this ensures that the cell is ready for DNA synthesis to occur. The second checkpoint occurs following the G2 phase after the necessary mitotic proteins are produced and the cell is once again subjected to regulatory mechanisms, to make sure that it is ready to enter the Mitotic (M) phase. This checkpoint is key to preventing the spread of damaged or mutated DNA from a single cell to daughters and granddaughter cells. DNA damage at this stage triggers the activation of the ATM/ATR kinases, which form a key part of the DNA damage response (DDR) signalling pathway, causing hundreds of proteins to be phosphorylated at Ser/Thr-Glu motifs and additional sites in an ATM- or ATR-dependent manner. ATT/ATR also activate a secondary wave of phosphorylation through their activation of Chk1, Chk2, and MK2 protein kinases. These mechanisms prevent the cell from passing into the M phase of the cycle until the damage to the DNA has been fixed. The final checkpoint in the cell cycle, the mitotic spindle checkpoint, occurs during the M phase of the cycle when all the chromosomes should have aligned at the mitotic plate and are thus considered to be under bipolar tension. This tension is sensed and is what initiates phase transition, resulting in the cell splitting into two daughter cells which then enter the G1 phase again. There is also a G0 phase (Gap 0) where the cell has left the cycle and stopped dividing. However, this only applies to those cells that no longer progress through the cell cycle temporarily or reversibly and thus are said to have entered a state of quiescence.

3.1.3 Stem Cell Differentiation

We are interested in the lineage commitment of key cell types along the haematopoietic lineage tree, understanding how to control cell lineage transition in cell cultures. The specific lineage tree for haematopoietic cells [71] is shown in Figure 3.2, including the different cell types as circles and the transition from one cell type to another illustrated by the arrows, starting with haematopoietic stem cells (HSC) on the far left, which become multipotent progenitor (MPP) cells. At this point, cells either become common myeloid progenitors (CMP) or common lymphoid progenitors (CLP). Following the CMP branch, we encounter another branching event, with cells differentiating and becoming either megakaryocyte-erythroid progenitor (MEP) or granulocyte-macrophage

progenitor (GMP) cells.

The section of the haematopoietic lineage tree we are interested in is the branch that includes CMP and MEP cells. CMP cells are capable of maturing into MEP and GMP cells, but we are interested in the maturation of CMPs to MEPs, as these eventually become mature red blood cells [19].

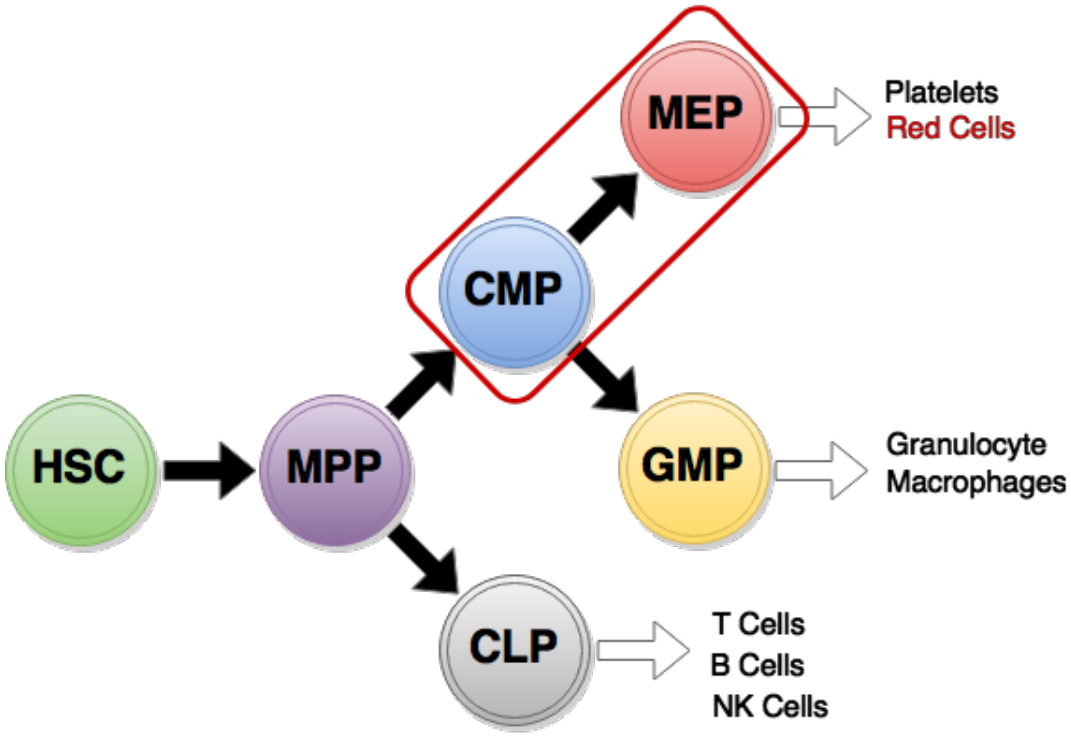


FIGURE 3.2. **Diagram illustrating the first two branching events along the haematopoietic lineage tree.** Haematopoietic Stem Cells (HSCs) are on the far left, becoming MultiPotent Progenitor (MPP) cells. At this point, cells can either differentiate into Common Myeloid Progenitors (CMP) or Common Lymphoid Progenitors (CLP) cells. Following along the CMP branch we encounter another branching event, with cells differentiating and becoming either Megakaryocyte-Erythroid Progenitor (MEP) or Granulocyte-Macrophage Progenitors (GMP). The segment of lineage tree we are interested in is outlined in red, tracking the progression from CMP to MEP.

Our focus in this Chapter is on trying to quantify the effects of erythropoietin (epo) on the CMP-MEP dynamics, as it is known to be fundamental to the control of both homeostatic and stress erythropoiesis (production of red blood cells under normal conditions as well as following trauma) [102]. Epo is vital in the later stages of erythroid proliferation: *in vivo*, epo is absolutely required for the CFU-E (Colony Forming Unit-Haematopoietic) stage, during which cells express greater levels of erythropoietin receptors and, as such, the presence of epo causes a larger proliferative response. The requirement for epo from the CFU-E stage onward has been shown,

with mutants defective for epo or epo receptor not producing mature red cells [181]. We considered the outlined segment of the lineage tree shown in Figure 3.2, and initialise our modelling and experiments from a population of CMP cells that have previously undergone the initial transitions of HSCs to CMP cells.

3.2 Experimental Methodology and Results

3.2.1 Methods

The cell-surface marker CD34 allows us to determine which cells are our initial population of CMP cells, with the presence of this marker ($CD34^+$) indicating that a cell is CMP. $CD34^+$ (CMP) cells were isolated from peripheral blood mononuclear cells (PBMCs) and were cultured in erythroid medium (EM) consisting of stem span (Stem Cell Technologies) + 10ng/mL stem cell factor (SCF) for the first 5 days of culture and + 50ng/mL for the subsequent 7 days of cell culture, 40ng/mL of insulin-like growth factor1 (IGF-1), 1ng/mL interleukin-3 (IL-3), $1\mu M$ dexamethasone and with or without 2 U/mL epo [177]. This protocol provides us with a starting population of CMP cells only, removing the possibility of alternative cells giving rise to our MEP population.

CMP and MEP cells were labelled using CFSE (carboxyfluorescein succinimidyl ester) which is a fluorescent cell staining dye. CFSE is cell permeable and covalently couples to intracellular molecules, including to intracellular lysine residues and other amine sources, via its succinimidyl group. The CFSE staining, due to its covalent coupling reaction, can be retained within cells for long periods of time. This means that when a cell divides, each daughter cell maintains a fluorescence that can correspond to the generation of the cells. Stained cells were put through a flow cytometer (MACSQuant) which fires lasers through the cells, exciting the stain, and causing the cells to emit light of varying wavelengths. The post-acquisition analysis involved gating on the scatter properties of the emitted light that corresponds to known surface markers', this was done with the software FlowJo v7.6.5 for the proliferation analysis and FloJo vX0.7 for all other analysis.

This analysis allowed us to capture lineage restriction events along the haematopoietic lineage tree. An example of the output of the flow cytometer data analysis, showing the gating on various surface markers, is shown in Figure 3.3. Following isolation, $CD34^+$ cells were 85% $CD34^+CD38^+IL3R\alpha^-CD45RA^-CD90^-CD45lo$ (left panel Figure 3.3) and, as such, could be defined as CMP cells [177]. CMP cells, cultured in serum-free medium with the addition of specific cytokines including SCF, IL-3 and epo, can mature into erythroid cells. To attempt to identify the more mature erythroid progenitors (MEPs) following lineage restriction from CMP cells, changes in the previously defined markers interleukin-3 receptor (IL-3R/CD123) and Fms-like tyrosine kinase 3 (Flt-3/CD135) were examined. The resulting expression of these markers on the total live population is shown in Figure 3.4.

The surface levels of both these markers increase over the first two days of the cell culture

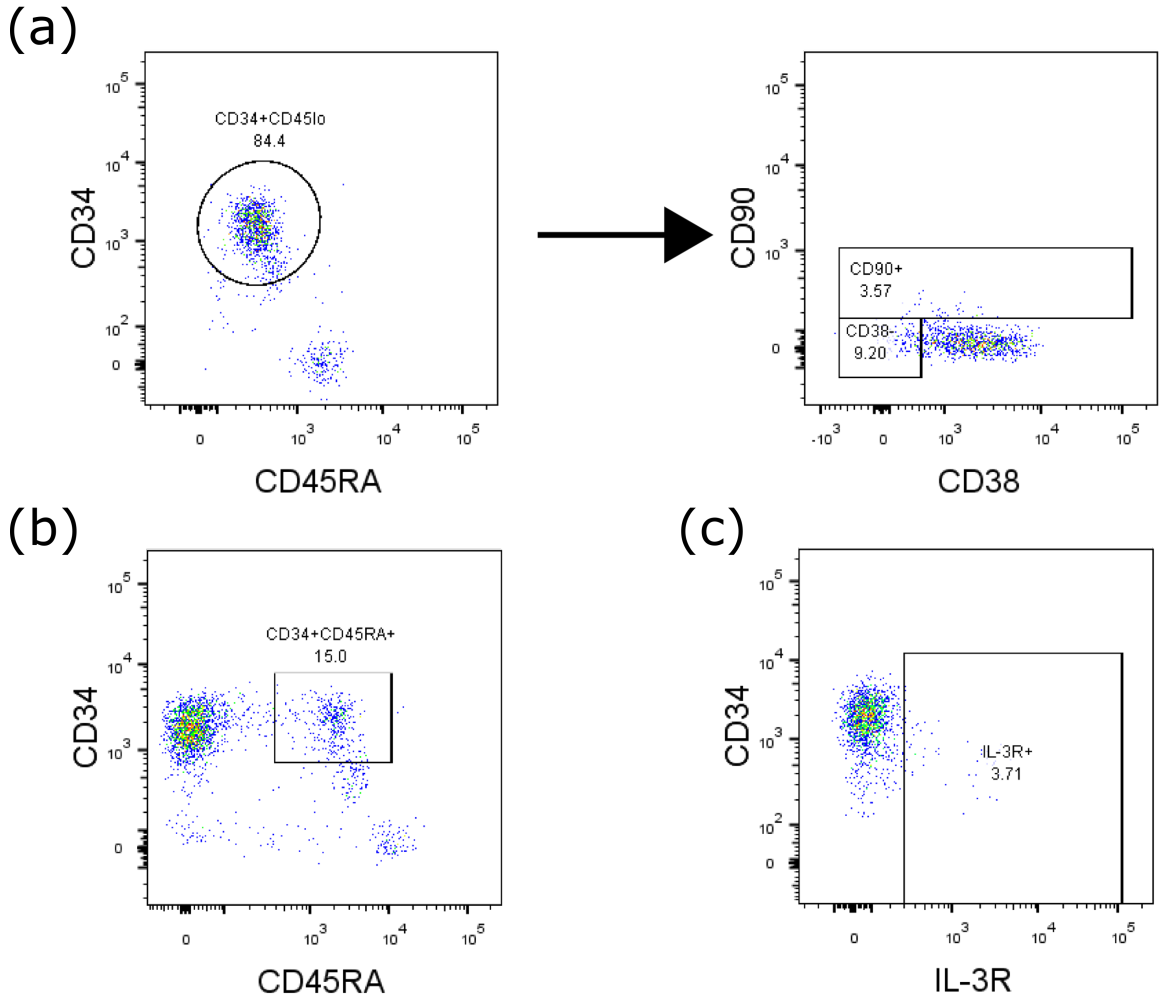


FIGURE 3.3. **Surface expression of key markers, including CD34, CD45, CD38, IL-3R(CD123), CD45RA and CD90, on peripheral blood (PB) CD34$^+$ cells.** Initial gating was done on CD34$^+$ and CD45lo (a, left), with subsequent gating on CD38 and CD90 (a, right). Analysis of IL-3R (b) and CD45RA (c) was completed on the total live population.

before then decreasing up to Day 4, making them unuseful as markers for restriction of CMP to MEP as neither of these markers are present in the more mature MEP cells. Alternatively, the erythroid marker CD36 was suggested to indicate the presence of the MEP cells, having originally been identified as a cell-surface marker of erythroblasts [52].

Following isolation, the CD34$^+$ cells do not express CD36, but eventually give rise to CD36$^+$ cells after a few days in culture. Figure 3.5 provides an example of how the expression of CD36 within the total live population of cells changes over four days, indicating that this marker is

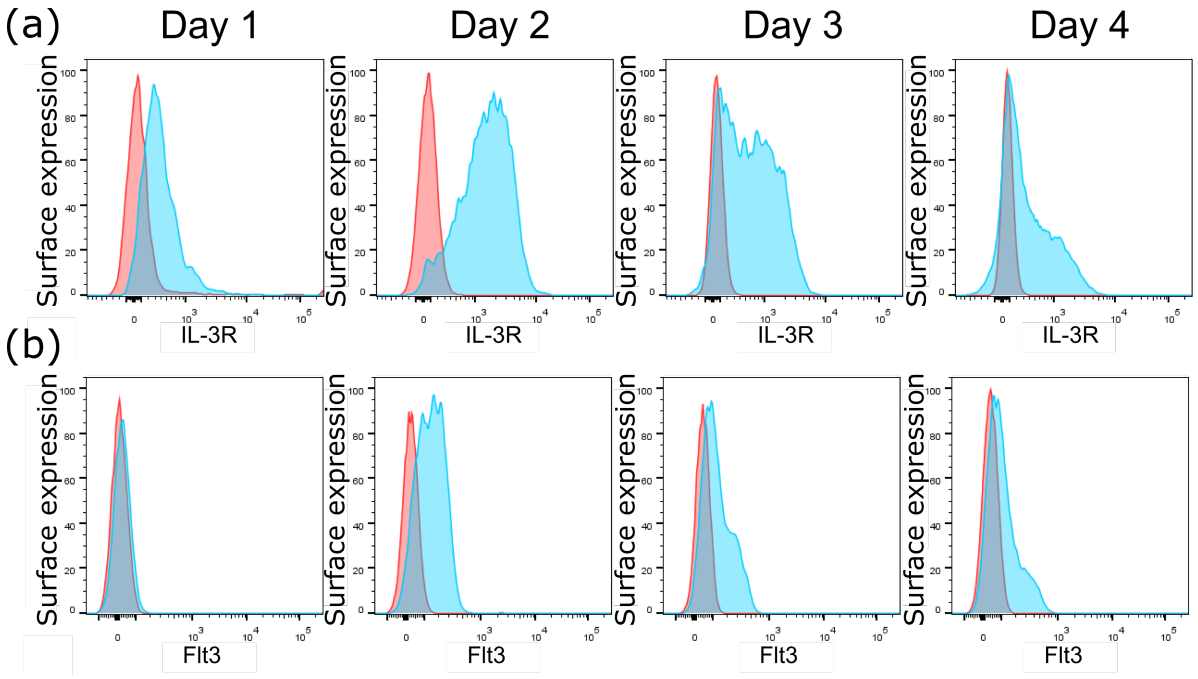


FIGURE 3.4. **Surface expression of (a) IL-3R (CD123) and (b) Flt3 on $CD34^+$ cells following culturing for 4 days.** The stained cells are indicated in blue and the isotype controls in orange. Expression of each surface marker increases in the population during the first three days of culture but then decreases over the subsequent day.

able to distinguish between the two cell subpopulations as indicated by the presence of two peaks from day 2 onwards. Although the dual peaks in Figure 3.5 indicate the presence of two distinct populations, in order to determine whether this surface expression of CD36 is able to mark the lineage restriction, $CD34^+$ cells were cultured for 2-4 days, after which the culture was separated into $CD34^+CD36^-$ and $CD34^+CD36^+$ populations, in order to determine the colony forming capacity of each of the two distinct populations. Figure 3.6 shows the resultant colonies, following sorting and further two weeks of culturing.

The $CD34^+CD36^-$ cells gave rise to both erythroid and myeloid colonies, as shown in Figure 3.6, indicating that they are the CMP population able to produce both CMP (through proliferation) and MEP cells (through differentiation). The $CD34^+CD36^+$ cells, on the other hand, gave rise to only erythroid colonies, which shows that this population has been lineage restricted to MEPs as these cells no longer produce the less specialised CMP cells. In addition, the erythroid colonies formed contained a small number of megakaryocytes, which is consistent with the assignment of $CD34^+CD36^+$ cells as bipotent MEPs. Having determined that the presence of the cell-surface marker CD36 can differentiate between the CMP and MEP cells within the cell culture, we can look at the specific cell counts for each of the two cell types.

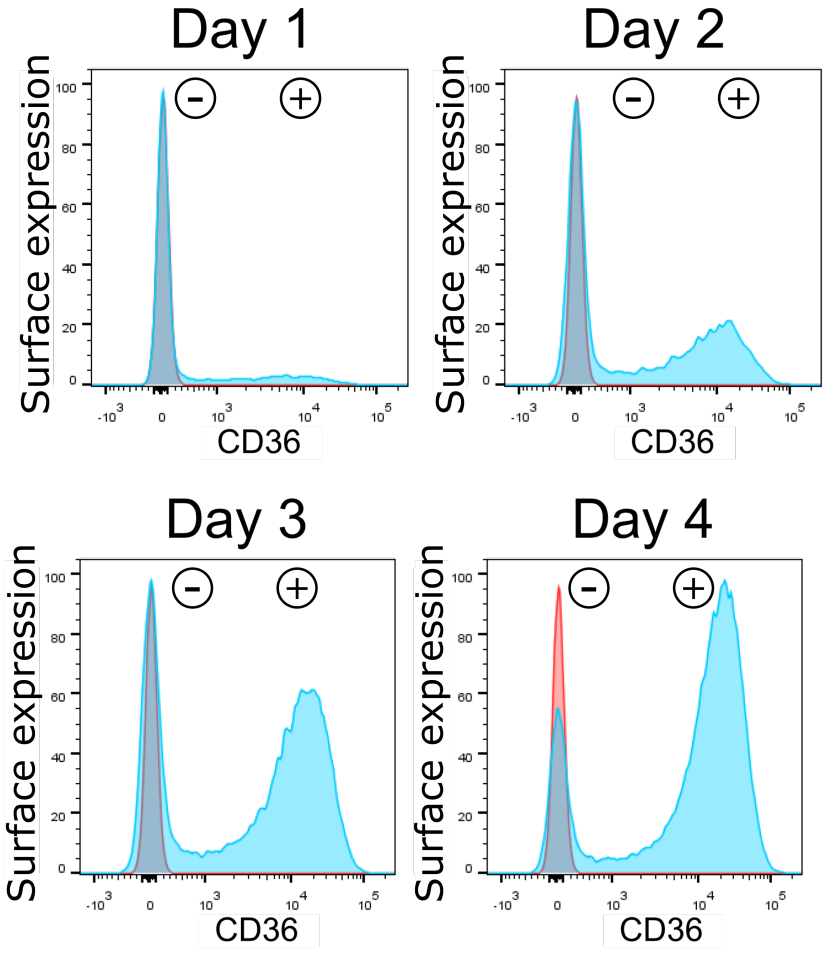


FIGURE 3.5. **Surface expression of CD36 on the total live population of cells, over four days of culture**. Surface expression demonstrates that $CD36^{-}$ and $CD36^{+}$ cells form two distinct populations which define the CMP and MEP populations. The (−) and (+) symbols denote the $CD36^{-}$ and $CD36^{+}$ populations. The stained cells are in blue and the isotype controls are in orange.

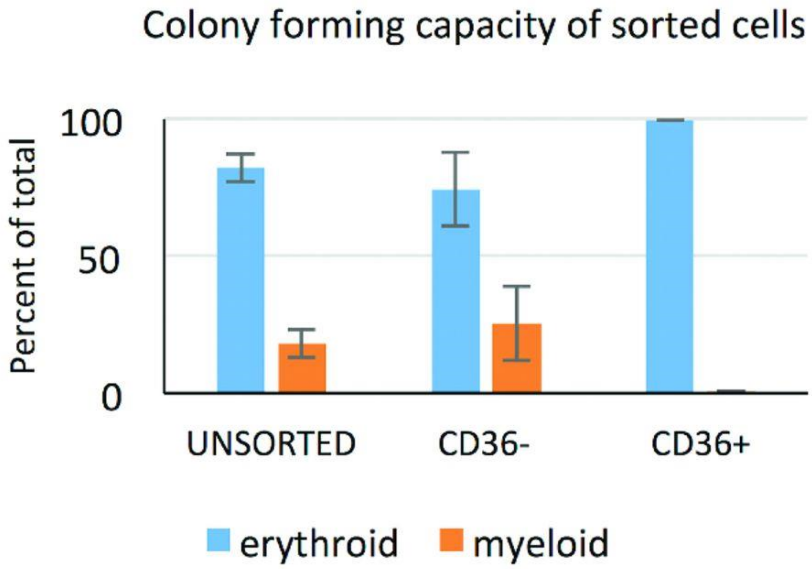


FIGURE 3.6. **Colony forming potential of the dual populations indicated by CD36 expression.** $CD34^+$ cells were cultured for four days, and sorted by FACS (fluorescence-activated cell sorting), a specialized type of flow cytometry, into $CD34^+CD36^-$ and $CD34^+CD36^+$ populations. These sorted populations were cultured for a further two weeks and the resultant colonies then counted.

3.2.2 Cell Culture Data

We then used the cell-surface marker CD36 to determine whether cells within the culture are CMP ($CD36^-$) or MEP ($CD36^+$). Figure 3.7 shows an example of results from the cell culture experiments over the first 4 days, following the culturing methods described in Section 3.2.1. The population was gated using the cell-surface markers CD34 and CD36 as previously mentioned, and the fluorescence of the resulting population was tracked. The x-axis indicates the wavelength of light emitted by the fluorescing cells: when a stained cell divides its daughter cells fluoresce less, which forms the distinct peaks present; this, in turn, means that each peak in the figure represents a different generation of cells.

The y-axis shows the number of cells that have a specific wavelength, thus enabling the quantification of the number of cells in each population and of the generation they reside in. Whilst this fluorescence data clearly shows the generational progression of each cell type, as a result of the proliferation of each population, we can convert fluorescence densities to individual generational cell counts. Doing so provides us with individual cell counts for each generation of the two populations over the course of the experiment, which can be readily used to fit a mathematical model. We will refer to generations of cell using the notation G(#), with # replaced by the generation of the cell. Considering the data in Figure 3.7, it is possible to see general behaviours, such as that after 24 hours in culture, all cells in both the CMP and MEP cultures

TABLE 3.1. **Mean (of three individual experiments) CMP (CD36-) cell count data from stem cell cultures (EDS1).** Data encompass six generations of cells (G0-G5). Data were collected at 8 hour intervals, starting from 10h (T10) after plating the culture until 90h (T90) after culture plating.

Time	G0	G1	G2	G3	G4	G5	Total
T10	7141.14	6.67	0	0	0	0	7147.81
T18	6334.55	126.18	0	0	0	0	6460.73
T26	5782.93	109.75	0	0	0	0	5892.67
T34	5111.43	586.83	38.03	0	0	0	5736.29
T42	4209.55	1580.16	222.76	29.04	0	0	6041.52
T50	3512.31	2176.47	696.88	68.59	0	0	6454.24
T58	2671.59	2310.79	1149.33	382.66	38.78	23.54	6576.69
T66	2071.77	2066.05	1558.58	1075.89	189.13	17.91	6979.34
T74	1717.34	1888.62	1762.33	1609.30	1092.07	46.59	8116.25
T82	1613.57	1543.72	2068.15	1872.64	1472.53	323.98	8968.87
T90	1402.56	1472.63	2064.48	2178.12	2603.37	1802.09	11890.16

remain in generation 0 (G(0)) and do not progress to G(1) until the second day of culture, which suggests that there is a lag phase for both cell types from the time of isolation to the initiation of proliferation, between days 1-2.

We have six sets of experimental data to consider. The first pair of datasets were collected using the standard experimental protocol (i.e. cells cultured in the presence of epo) and will be referred to as Experimental Dataset 1 (EDS1) and Experimental Dataset 2 (EDS2). These datasets were used to fit and validate our mathematical model, see Sections 3.3.3 and 3.5. The second pair of datasets were collected using competing experimental protocols, with one set, Experimental Dataset 3 (EDS3), culturing cells in the presence of epo and the other, Experimental Dataset 4 (EDS4), culturing cells in the absence of epo, to provide directly comparable data and were used to elucidate the effects of epo on cell proliferation and differentiation (Section 3.6). The final pair of datasets were collected to provide a comparison with our model predictions when extrapolating beyond the initial culturing stage (1-4 days), with cells cultured both with/without epo. They will be referred to as Experimental Dataset 5 (EDS5) and Experimental Dataset 6 (EDS6).

Tables 3.1 and 3.2 contain the experimental data from EDS1, showing how the mean cell counts for CMP (CD36-) and MEP (CD36+) cells, respectively, over the course of the culturing experiment. The individual counts were extracted from fluorescence data collected by flow cytometry; Figure 3.7 presents an example of the data collected from a single experiment. The culturing experiments were performed in triplicate in order to provide a more reliable data set, making it possible to remove erroneous outlying values that can arise due to small changes in the culture conditions. The mean total cell counts (shown the rightmost column in Table 3.1)

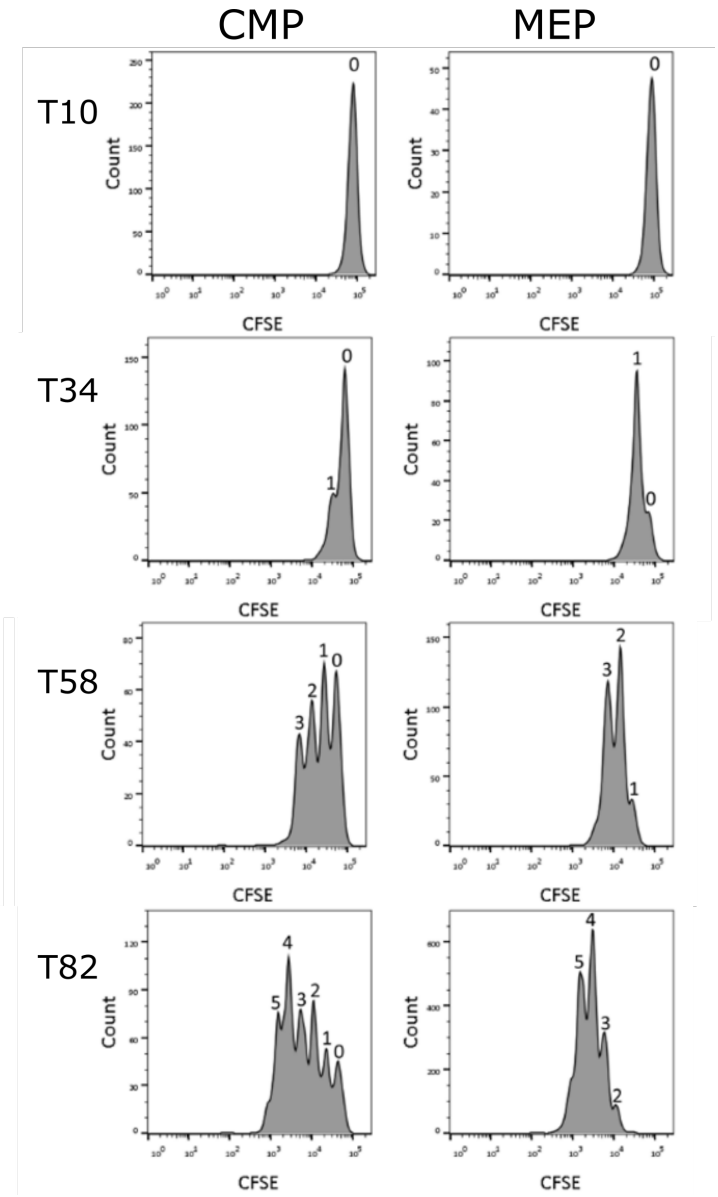


FIGURE 3.7. **CFSE data collected from haematopoietic stem cell culture experiments.** Figure showing the CFSE (fluorescent dye) data collected from a single experiment using flow cytometry, tracking the progression of CMP and MEP cell populations over 4 days. The emergence of peaks is indicative of cellular proliferation, with newly divided cells maintaining a reduced fluorescence than their mother cells. This data can be translated into individual counts for each cell type in each generation.

TABLE 3.2. **Mean (of three individual experiments) MEP (CD36+) cell count data from stem cell cultures (EDS1).** Data encompass six generations of cells (G0-G5). Data were collected at 8 hour intervals, starting from 10h (T10) after plating the culture until 90h (T90) after culture plating.

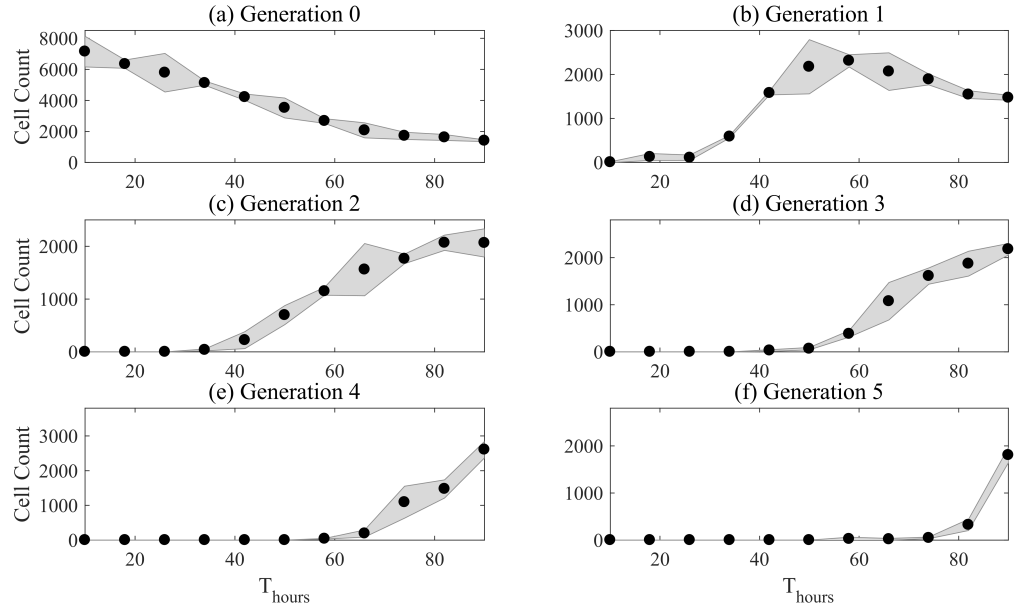
Time	G0	G1	G2	G3	G4	G5	Total
T10	516.32	0.00	0.00	0.00	0.00	0.00	516.32
T18	1523.72	87.83	0.00	0.00	0.00	0.00	1611.55
T26	1859.66	133.62	5.92	0.0 0	0.00	0.00	1999.19
T34	1443.91	1209.83	150.50	19.70	0.00	0.00	2823.98
T42	682.76	2625.14	357.38	39.87	0.00	0.00	3705.15
T50	300.07	2968.23	2296.30	280.23	18.17	0.00	5863.01
T58	76.82	2333.79	4579.71	1681.98	216.50	10.58	8899.38
T66	48.63	1253.80	4982.65	5269.89	1143.72	143.95	12849.99
T74	43.01	961.47	4922.79	8261.08	5167.94	897.92	20273.75
T82	0.00	557.61	3590.13	8428.38	10055.94	4406.02	27551.13
T90	0.00	252.91	2406.31	8423.00	15237.82	13148.57	43478.78

indicate that the total population of CMP cells does not increase significantly when compared to the MEP cell population (total in Table 3.2), with a fold change vs the first time point of 2.35, calculated for the CMP population over the course of 90 hours, and a fold change of 60.58 for the MEP population over the same period.

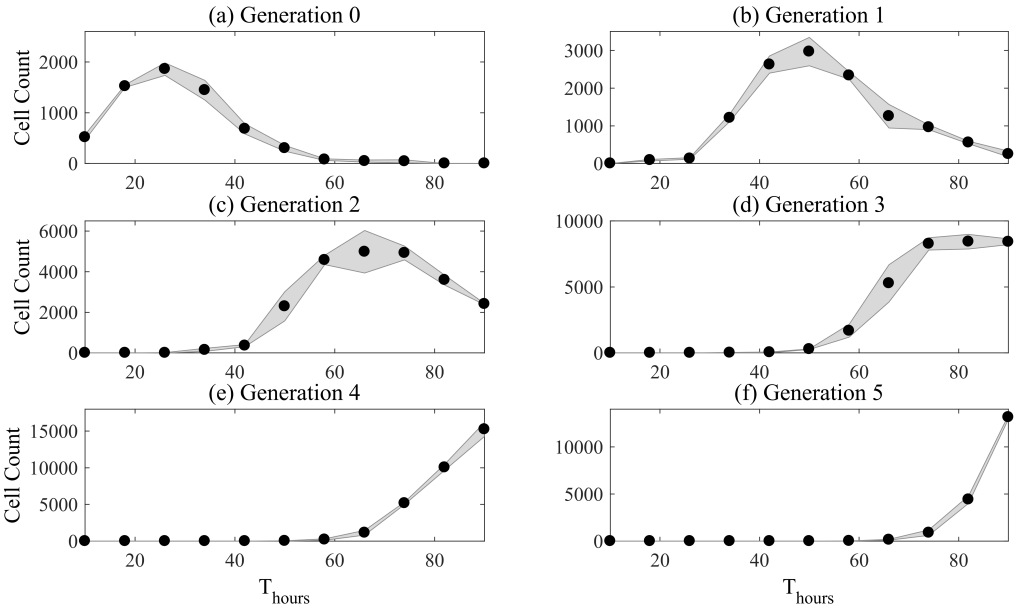
Comparing the data presented in Tables 3.1 and 3.2 it appears at first glance that the proliferative capacity of the MEP cells is significantly larger than that of the CMP cells, suggesting that fast differentiation of CMP to MEP would be able to achieve such population amplification. However, it also appears that there is a delay between the appearance of a specific generation of CMP and MEP and the appearance of a subsequent generation. This delay is clearer having plotted the cell counts for each generation of the two populations, as shown in Figure 3.8(a) for the CMP cells and Figure 3.8(b) for the MEP cell population: CMP cells arise in G(1) after 26 hours and subsequently in G(2) after 42 hours, whilst there is a similar delay in the MEP proliferation.

At any point, a CMP cell has the capacity to either travel through its cell cycle and divide to amplify the CMP population itself or to differentiate and become an MEP cell. It is not clear from the experimental data cell counts alone which of these processes is more important to maintaining a high amplification system for these stem cells, nor whether the differentiation of CMP cells to MEP cells is a constant process or whether it has any temporal confinements. In order to try and capture the proliferative and differentiative dynamics of the stem cell population, we want to fit parameters that quantify these two processes, providing some evidence as to the temporal focuses of the proliferating and differentiating populations.

An additional complexity is the possible effects of cell death (apoptosis) on the cell culture. To



[a]



[b]

FIGURE 3.8. **Generational data for the CMP (a) and MEP (b) populations within the cell culture.** Black dots represent mean cell counts from three experiments, shaded region indicates the standard deviation. Sampling time is 8h and total experiment time is 90h. Data shown in Tables 3.1 for (a) and 3.2 for (b).

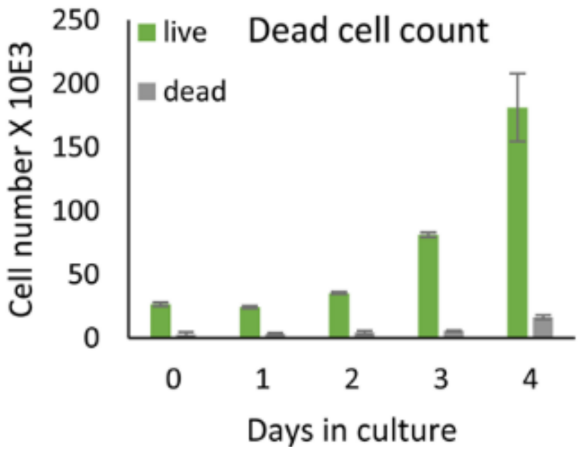


FIGURE 3.9. **Experimental data showing the influence of cell death in the culturing of HHSCs over 4 days of culture.** The green bars show the number of live cells in the culture, the grey bars show the number of dead cells. The error bars represent the standard error calculated across three separate experiments. The number of dead cells is significantly smaller and insignificant compared to the number of living cells (approximately 8%) and does not appear to increase over the course of culture.

investigate whether this process plays a significant role in the culturing of HHSCs we tracked the number of cells undergoing apoptosis during the experiment. As shown in Figure 3.9, the number of living cells increases exponentially over the course of the experiments, however the number of cells undergoing apoptosis appears to remain consistent, with approximately $5\% \pm 1\%$ of the population undergoing apoptosis at day 0 and $6\% \pm 1\%$ after 4 days. Thus, in what follows, cell death was not modelled.

3.3 Modelling Haematopoietic Stem Cell Populations

3.3.1 Random Birth-Death Model

To allow us to quantify the dynamic behaviours in the cell counts of the two subpopulations we can consider a number of possible modelling approaches that describe these dynamic populations, with different possible approaches detailed in Chapter 2.1. We initially attempted to apply the random birth-death (RBD) modelling approach, detailed in Chapter 2 Section 2.1.1, to our HHSC culture. As previously described, the random birth-death model assumes fixed homogeneous division and death rates α and β.

Figure 3.10 presents the resultant fitting of a RBD model to our experimental data. The fitted birth and death rates are $\alpha = 0.05$ and $\beta = 0.04$ respectively. It is clear that the initial population of CMP cells in G(0) begins maturing immediately, and as such so do subsequent generations

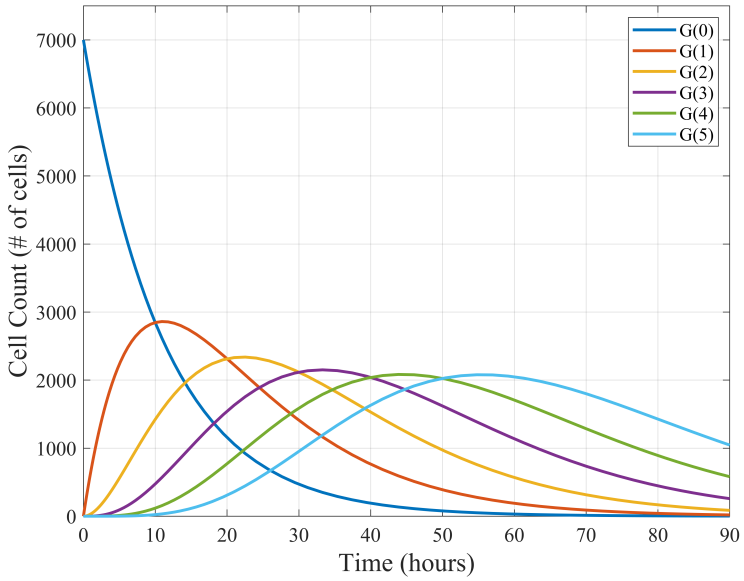


FIGURE 3.10. **CMP generational cell counts from random birth-death model.** Changing cell counts governed by Equations 2.1 and 2.2 in Chapter 2. Birth and death rates, $\alpha = 0.05$ and $\beta = 0.04$, fitted to experimental data (EDS1).

of the cells in our modelled population. The issue we found with this modelling approach is that, when applied to the experimental data (Figure 3.8), there is a rapid onset of proliferation which does not accurately represent the lag we observed over the first 24 hours of cell culture (Figure 3.8 and Tables 3.1 and 3.2). In addition, using a RBD model we are unable to impose a minimum cell-cycle time, something that we observed in the experimental data. Due to the inability of this model to account for the delays seen in the onset of cell proliferation, as well as the delay in maturation of cells to subsequent generations as required, it is not a suitable model for replicating the dynamics of HHSC culture.

3.3.2 Adapted Smith-Martin Model

The model we have used is an adaptation of the so-called Smith-Martin (SM) model [155], which has been used successfully to model and predict cell counts from CFSE-based cytometry data, as well as to derive biologically significant parameters [7, 129, 155]. The original Smith-Martin model splits the cell-cycle of a single cell into two distinct compartments “a” and “b”, the former representing the G1 phase of the cell cycle and the latter representing the combined S, G2 and M phases (Figure 3.1).

We adapted the Smith-Martin model to include more than one cell type. This allows us to model the maturation of CMP to MEP, as shown schematically in Figure 3.11. This transition from the “a” compartment of CMP to the “a” compartment of MEP can be described by a rate

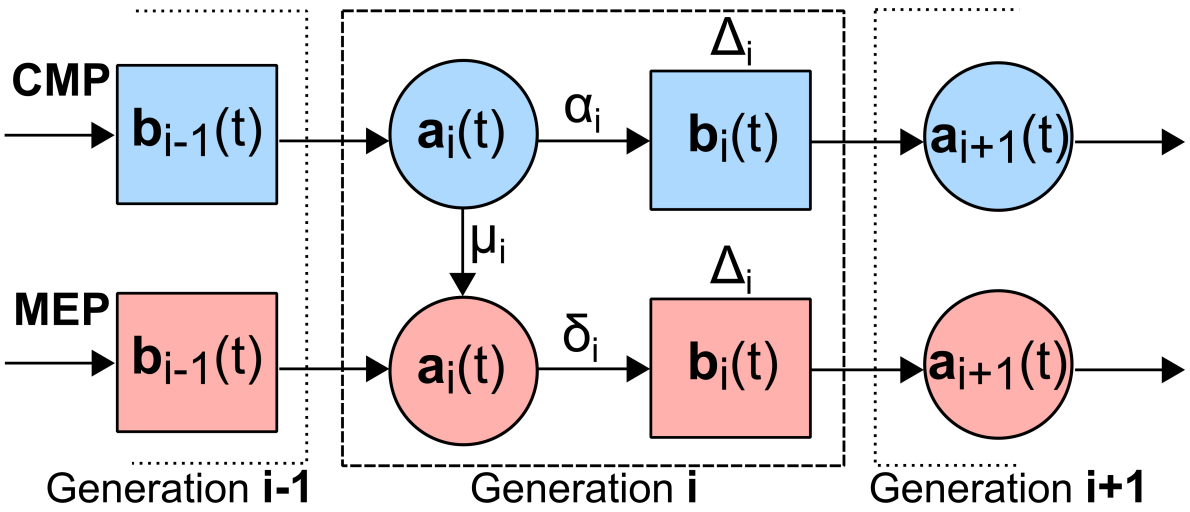


FIGURE 3.11. **Diagram of the DDE (adapted Smith-Martin) model of the cell cycle.** Mathematical model of the erythroid cell culture system. The cell cycle is modelled as a 2 stage process, composed of compartments "a" approximating the G1 phase of the cell cycle, and "b" approximating the S/G2/M phases. CMP and MEP both originate in the "a" compartment. Their transition is governed by a stochastic process, with rates α_i and δ_i for generation i CMP and MEP cells respectively to transition to the "b" compartment, and a rate μ_i of maturing (transitioning) from CMP into same-generation MEP cells. Cells remain in the "b" compartment for a fixed time (Δ_i and Γ_i for CMP and MEP respectively) before they divide and then re-enter the "a" compartment as a next-generation cell.

μ_i for generation i CMP maturing to generation i MEP. A cell that has transitioned into the "b" compartment remains there for a fixed time (Δ_i and Γ_i for CMP and MEP respectively) before it divides and then re-enters the "a" compartment as a next-generation cell. A cell is initialised in the "a" compartment, with its transition to the "b" compartment governed by a fixed probability. The latter probabilities are described by the parameters α_i and δ_i for generation i CMP and MEP cells respectively, for transitioning from the "a" to the "b" compartments (Figure 3.11). The stochastic behaviour of the "a" compartment and the fixed time delay of the "b" compartment relate directly to observed biological dynamics [155]. The ability to fit parameters linked to proliferation and differentiation, whilst incorporating the delay alluded to in Figure 3.8, makes the adapted Smith-Martin model a good formalism for our biological system.

The original Smith-Martin model does include further parameters that deal with cell apoptosis (cell death). However, as aforementioned, we have observed that during this initial early phase of cell production cell apoptosis is relatively constant (Figure 3.9) and can, therefore, be neglected in the model. This allows us to make the model simpler, improving computational efficiency without sacrificing accuracy.

In summary, in order to simplify the model whilst encompassing the key dynamics that relate to lineage restriction, we made the following assumptions based on the experimental data: (1) death rate plays an insignificant role over the time course (Figure 3.9); (2) cell maturation initiates prior to the onset of CMP and MEP proliferation (Figure 3.8) and, thus, does not require transition through the cell cycle. The CMP and MEP populations in each generation evolve according to the following set of coupled delay differential equations (DDEs)

$$\dot{C}_0(t) = -\mu_0(t)C_0(t) - \alpha_0(t)C_0(t), \tag{3.1}$$

$$\dot{C}_i(t) = 2\alpha_{i-1}C_{i-1}(t-\Delta_i) - \mu_i(t)C_i(t) - \alpha_i(t)C_i(t), \tag{3.2}$$

$$\dot{M}_0(t) = \mu_0(t)C_0(t) - \delta_0(t)M_0(t), \tag{3.3}$$

$$\dot{M}_i(t) = 2\delta_{i-1}M_{i-1}(t-\Gamma_i) + \mu_i(t)M_i(t) - \delta_i(t)M_i(t), \tag{3.4}$$

where the number of cells in any generation of CMP (C) and MEP (M) are dependent on the cell counts in the previous generation, and an earlier time. To account for the observation that maturation and proliferation, for the first generation, are not synchronous (maturation largely occurs prior to the initiation of proliferation as shown in Figure 3.8 and Tables 3.1 and 3.2), we set the first generation proliferation and maturation rates to be time-dependent, in the form of generalised logistic functions. The latter model the time-dependence of the proliferation and maturation rates as they allow us to investigate a range of time-dependence, from a rapid rate switch (turning proliferation/maturation on/off immediately), to a slower ramping up or down of the rates. The logistic functions were fitted to experimental data. We defined the logistic functions for the proliferation rates of both CMP and MEP, and the maturation rate as follows

$$\alpha_0(t) = \frac{P_\alpha}{(1+Q_\alpha e^{-R_\alpha(t-\tau_C)})^2}, \tag{3.5}$$

$$\delta_0(t) = \frac{P_\delta}{(1+Q_\delta e^{-R_\delta(t-\tau_M)})^2}, \tag{3.6}$$

$$\mu_0(t) = \mu_I + \frac{\mu_F - \mu_I}{(1+Q_\mu e^{-R_\mu(t-\tau_C)})^2}, \tag{3.7}$$

where $\alpha_0(t)$ (Equation 3.5) is the proliferation rate of generation 0 (G(0)) of the CMP population, $\delta_0(t)$ (Equation 3.6) is the proliferation rate of G(0) of MEP cells, and $\mu_0(t)$ (Equation 3.7) is the maturation rate of the G(0) CMP cells to MEP cells.

In Equation 3.5, the time-dependent G(0) CMP proliferation function (P_α) is the maximal proliferation rate of G(0) CMP, Q_α is a scaling term that provides micro adjustments to the start of the logistic function, R_α is the logistic function growth rate governing the steepness of the logistic function, and τ_c sets the start time for the logistic function for CMP.

In Equation 3.6, governing the proliferation of G(0) MEP, P_δ is the maximal proliferation rate of G(0) CMP, Q_δ is a scaling term that provides micro adjustments to the start of the logistic function, R_δ is the logistic function growth rate governing the steepness of the logistic function, and τ_m is the onset time for MEP proliferation.

The final logistic function in Equation 3.7, which governs the maturation of CMP to MEP, differs from the previous functions in Equations 3.5 and 3.6 in that it includes parameters μ_F and μ_I which are the final and initial rates of maturation, respectively. These are included as, unlike the proliferation of G(0) CMP and MEP cells (which begins at zero and plateaus at a maximal value), the maturation rate starts at a maximal value and decreases to a minimal value that is greater than zero, as maturation does not necessarily cease entirely. As in Equations 3.5 and 3.6, Q_μ is a scaling term that provides micro adjustments to the start of the logistic function, R_μ is the growth rate affecting curve steepness and τ_c is the onset time for CMP to MEP maturation.

The parameters for these logistic functions are fitted alongside the parameters for the adapted Smith-Martin model (Equations 3.1 - 3.4) in order to minimise the root mean square error (RMSE) to the experimental data, defined as sample standard deviation of the differences between simulated values and training data; the details are provided in Section 3.4, and the resultant curves are shown in Figure 3.15. Considering the whole system, with six generations of CMP and MEP cells, we have 39 parameters to fit. This large number of possible parameters to vary could lead to an over-fitting of the model to the available experimental data. Thus, we considered possible simplifications to reduce the number of unknown parameters in the model. In the following section, we detail the methods for fitting the model parameters.

3.3.3 Model Fitting and Parameter Optimisation Algorithms

Assuming all parameters are generationally-dependent and as such require individual rates for proliferation in each generation of CMP and MEP, individual maturation rates for each generation of CMP, an individual parameters for each logistic function version of proliferation of G(0) CMP and MEP, and G(0) CMP maturation, we have 39 parameters to optimise. The fitting of the model required setting an objective function, which in our case is the average root mean squared error (RMSE) between the generational cell counts of CMP and MEP cells collected experimentally and the generational cell counts predicted by our adapted Smith-Martin model. The aim is to optimise the unknown parameters such that the average RMSE between the generational counts

predicted by our model and the experimental data is minimal, whilst the parameters still fall within a realistic range.

For parameter fitting we relied on a hybrid simulated annealing algorithm [166, 167], which involved a combination of simulated annealing and pattern searching. The simulated annealing algorithm was used due to the large parameter space, with simulated annealing being an efficient method to scan it. The simulated annealing algorithm is inspired by the process of annealing in metalwork, a process involving heating and cooling a material in order to alter its properties due to changes in internal structuring. As metal is cooled, it fixes its new internal structure and retains its newly obtained properties. As such, in the simulated annealing algorithm, a 'temperature' variable is maintained (which simulates the 'heating' process). Simulated annealing works by varying the model's input parameters such that it achieves a better fit to the objective function, but it maintains an ability to accept and briefly move in the direction of a worse solution. This occurs with changing frequency depending on the 'temperature' variable: whilst the 'temperature' is high, the algorithm accepts, with higher frequency, solutions which are less optimal than our current solution. This is used to avoid getting trapped in local minima, possibly found in previous iterations. As the 'temperature' is reduced so is the frequency of accepting less optimal solutions, gradually allowing the algorithm to converge towards an area in the parameter space where an optimal solution can be found.

In combination with the simulated annealing algorithm, we also used the pattern search algorithm. The latter algorithm differs from simulated annealing in that it does not accept worse solutions to the objective function at any stage of the algorithm. The pattern search algorithm finds a sequence of points in the solution space that approaches an optimal point.

The hybrid method we used takes the final solution from the simulated annealing algorithm and uses a mesh to find additional solutions in a pattern in the surrounding solution space. It then polls these points, checking them in order, to find a point which moves towards an optimal solution. It then updates this point in the solution space and repeats by finding new solutions in the surround solution space and moving towards the optimal solution. Thus, pattern search was used secondary to the simulated annealing algorithm, to try and move as close to the optimal solutions as possible.

The final optimisation method that we used was the fmincon routine in MATLAB [167]. This routine uses a trust-region approach to optimisation, which approximates the objective function f with a simpler function q which reasonably reflects the behaviour of the objective function in a neighbourhood N around the point which is the trust-region. The approximation q is defined by the first two terms of a Taylor approximation to F at the solution point x, with a spherical neighbourhood N. The trust-region problem is typically stated as

$$min\{\frac{1}{2}s^T Hs + s^T g \quad \text{such that} \quad \|Ds\| \leq \Delta\}, \tag{3.8}$$

where g is the gradient of f at the current point x, H is the Hessian matrix (the symmetric

matrix of second derivatives), D is the diagonal scaling matrix, Δ is a positive scalar, and $\|\quad\|$ is the 2-norm. The fmincon optimising routine restricts the trust-region subproblem to a two-dimensional subspace S, determined with the aid of a preconditioned conjugate process. It defines S as the linear space spanned by s_1 and s_2, where s_1 is in the direction of the gradient g and s_2 is either an approximate Newton direction and as such a solution to

$$H \cdot s_2 = -g, \tag{3.9}$$

or a direction of negative curvature,

$$s_2^T \cdot s_2 < 0. \tag{3.10}$$

This choice of S is included to force global convergence (using the steepest descent direction or negative curvature direction) and to achieve the fastest local convergence (using the Newton step if it exists). The algorithm works by formulating the two-dimensional trust-region subproblem. It then solves Equation 3.8 to determine the trial step s. If $f(x+s) < f(x)$ then $x = x+s$. Following this, it then adjusts the value of Δ and repeats until convergence. The trust-region dimension Δ is decreased if the trial step is not accepted.

Using these optimisation methods we can consider the whole model case and analyse whether the model can be reduced to simplify the process of achieving the most optimal fit to the experimental data. The options specified for the parameter identification were left at the default values in both optimtool and fmincon.

3.3.4 Model Reduction

The initial model setup, as presented in Figure 3.11, assumed a general case of full system heterogeneity (i.e. all 39 parameters, covering the proliferation rates of CMP and MEP cells, as well as the maturation of CMP to MEP, the proliferation delays and the logistic function parameters, are generation-dependent). Fitting all of the 39 parameters to the experimental data using the hybrid simulated annealing algorithm described in Section 3.3.3 provided insight as to the general system behaviour and allowed us to make simplifications.

Table 3.3.3 shows the fitted parameters for the whole 39 parameter model (as a result of assuming full system heterogeneity). These best-fit parameters indicated that the system is homogeneous (with each generation proliferating at the same rate) for MEP proliferation and heterogeneous (with each generation proliferating with different rates) for CMP proliferation. The delays between the “b” and “a” compartments for all generations of CMP are also homogeneous, with each generation of CMP cells being delayed in subsequent generations by the same fixed delay, whilst the delays for generation 4 and onwards for the MEP cell subpopulation also appearing to be homogeneous whilst the first three generations required their own fitted delays.

TABLE 3.3. Optimised parameters for the full DDE (adapted Smith-Martin) model of cell proliferation and differentiation, incorporating all 39 parameter values.

Parameter	Value
CMP proliferation, P_α	0.0183
α_1	0.0541
α_2	0.1067
α_3	0.1789
α_4	0.2088
α_5	0.1951
Initial maturation, μ_I	0.0165
Final maturation, μ_F	0.00393
μ_1	0.00404
μ_2	0.00403
μ_3	0.00405
μ_4	0.00402
μ_5	0.00404
MEP proliferation, P_δ	0.126
δ_1	0.1218
δ_2	0.1154
δ_3	0.1211
δ_4	0.1217
δ_5	0.1098
CMP initiation delay, τ_c	30.39 hours
MEP initiation delay, τ_m	27.71 hours
CMP proliferation delay, Δ_0	9.791 hours
Δ_1	9.803 hours
Δ_2	9.805 hours
Δ_3	9.762 hours
Δ_4	9.806 hours
Δ_5	9.805 hours
MEP proliferation delay, Γ_0	14.55 hours
Γ_1	9.817 hours
Γ_2	8.017 hours
Γ_3	8.021 hours
Γ_4	8.002 hours
Γ_5	8.002 hours
Logistic function parameters, Q_α	0.58
R_α	0.44
Q_δ	0.56
R_δ	0.43
Q_μ	0.618
R_μ	0.065

As the cell culture system did not appear to be sensitive to these parameters, with a best fit achieved without varying them, we can instead allocate single parameters for these rates.

So in what follows we use a single delay for CMP and three delays for MEP (defining $\Gamma_{(i\geq 4)} = \Gamma_3$). Also, the fitted maturation rates for later generations (G>0) were significantly smaller (10%) indicating that the influence of maturation on the system dynamics arose primarily from G(0). To account for the small effect of maturation in later generations, we defined $\mu_{(i\leq 1)}$=μ_F as the final maturation rate of G(0). The parameters that govern the dynamics of the generalised logistic equations (Q_α,R_α,Q_δ,R_δ,Q_μ,R_μ) were determined using EDS1 (CFSE tracking of CMP and MEP generational cell counts in Figure 3.7) and then remain fixed throughout. These simplifications reduced the system to 20 parameters to fit (not including the fixed logistic functions): 12 proliferation rates ($\alpha_0 - \alpha_5$, $\delta_0 - \delta 5$), 2 maturation rates (μ_I, μ_F), 4 compartmental delays (Δ, $\Gamma_{(1-3))}$ and 2 proliferation initiation delays (τ_C, τ_M).

3.4 Model Fitting and Results

We then fitted the simplified DDE (adapted Smith-Martin) model to the experimental data collected, using the objective function and fitting algorithms described in Section 3.3.3. The training data (EDS1) consists of CFSE tracking data from which we extracted CMP and MEP generational cell counts (Figure 3.7). We initially ran a hybrid simulated annealing algorithm [166] (*simulannealbnd* solver in the MATLAB optimtool [166] routine, combined with pattern search performed at the end of the simulated annealing) approximately 30 times in a large range of initial values and bounds. The bounds were set at [0 1] for all rates, initial estimation $\pm$ 2 hours for delays and initial estimations $\pm$ 4 hours for τ_C and τ_M, using the experimental data to suggest initial parameter estimations where possible. Best fittings (in terms of lowest RMSE) were selected to further refine the parameter estimations and optimise the bounds, via the MATLAB fmincon routine.

The results from the optimisation routines for parameter identification in the reduced model are shown in Table 3.4, with the resultant rates (divisions per hour) and delays (hours) obtained from the optimisation algorithm. We observed a very good quantitative agreement in cell count for all generations (Figures 3.12 (a) and (b)). To calculate the absolute lower and upper bounds of experimental error, we calculated the standard deviation across generational cell count averages (shaded regions in Figure 3.12). The absolute lower bound of experimental data was 94.74 cells (1.37%) and the upper bound was 259.42 cells (3.69%) for CMP, and 100.54 cells (0.95%) and 305.98 cells (3.22%) for MEP. The RMSE error for the total cell counts, found in the model fitting in Figure 3.13 was 101.02 cells (1.38%) for CMP and 138.29 cells (1.85%) for MEP.

The identified parameters (Table 3.4) show that the average MEP proliferation rate (δ) is 0.12 divisions per hour and this rate does not vary significantly from the onset of proliferation to the end of the culture experiment. The rate of proliferation of CMP cells is initially approximately

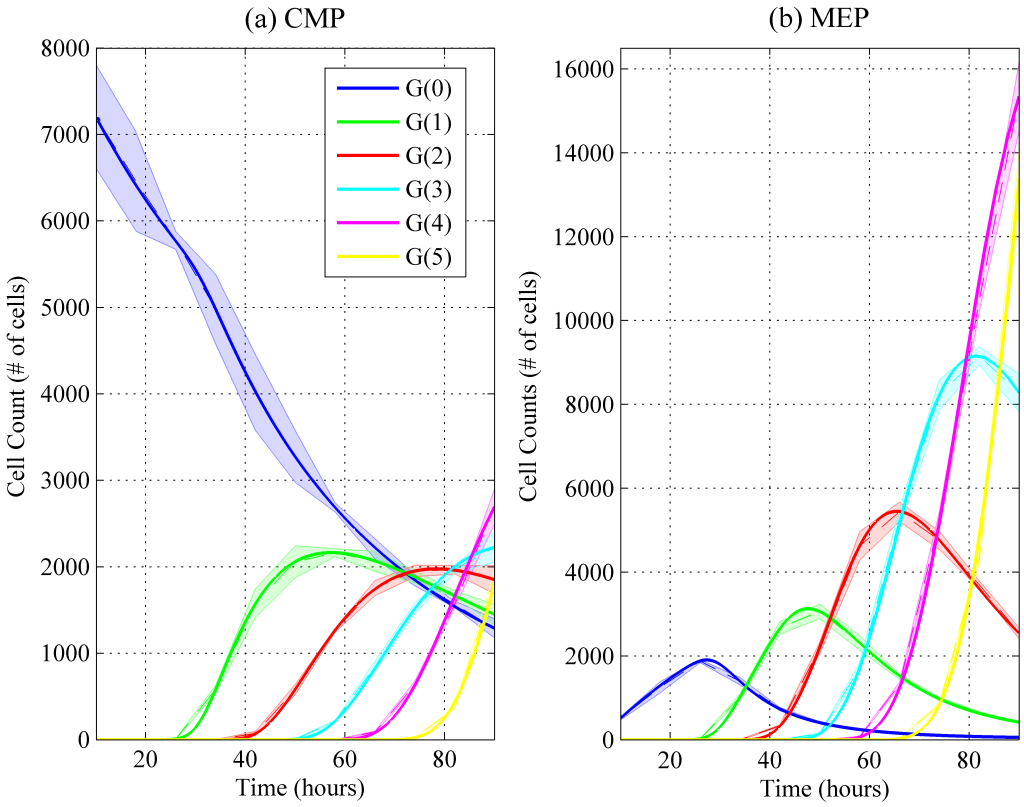


FIGURE 3.12. **Generational cell count results from fitted DDE (adapted Smith-Martin model).** Fitted model simulations provide proliferation and maturation parameters and recapitulate population dynamics: fitted model simulations (solid lines) of CMP (left panel) and MEP (right panel) cell numbers in each generation. Experimental data are dashed lines, +/-standard error (shaded regions).

10% of the proliferation rate of MEP and increases with each generation such that CMP transit from G(5) to G(6) at approximately twice the average MEP rate, as shown graphically in Figure 3.14. The maturation rate for G(0) CMP cells decreases over time, as shown in Figure 3.15(c), from 0.016 to 0.004 transitions per hour. Importantly, the maturation rate, even at the beginning of the culture when it is at its highest, is an order of magnitude smaller than the MEP proliferation rate. The relative contributions of proliferation and maturation to the overall population dynamics at different stages in the culture inducated by the identified parameters: while maturation plays a crucial role at the beginning, proliferation is the main effector of later culture stages.

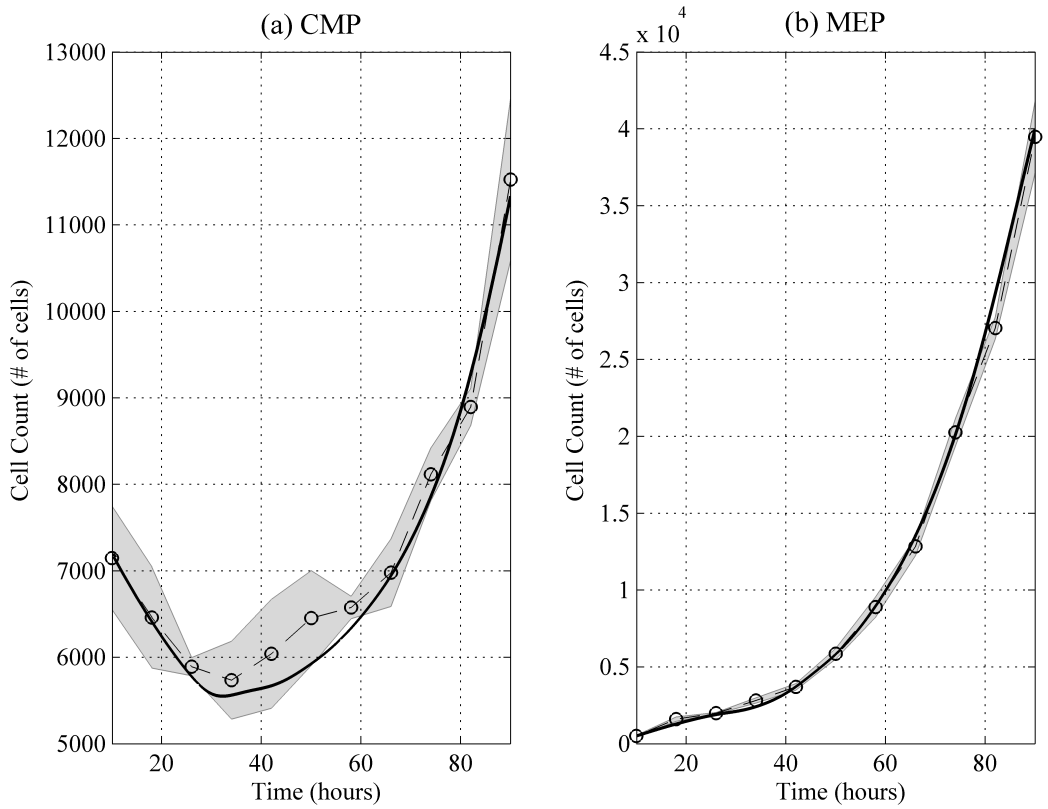


FIGURE 3.13. **Total cell count results from fitted DDE (adapted Smith-Martin model).** Fitted model simulations provide proliferation and maturation parameters and recapitulate population dynamics: Model simulation of total CMP (left panel) and MEP (right panel) cell numbers using model derived parameters (solid lines) together with experimental data (dashed lines) +/- standard error (shaded regions).

3.5 Model Validation, Robustness and Sensitivity Analysis

In order to validate the model fitted on the first dataset (Experimental dataset 1 (EDS1)), we used an additional set of experimental data (Experimental dataset 2 (EDS2)) which was comprised of an additional three CMP/MEP cultures data.

Figure 3.16 shows validation results: using the parameters fitted to EDS1 we used the adapted Smith-Martin model to predict the population dynamics of EDS2, using the initial cell counts for the CMP and MEP cells. As shown by the predicted dynamics (dashed red line), the model shows qualitatively similar dynamics but underestimates cell counts in G(2) onwards in the CMP population. The prediction for the MEP counts is within the standard deviation of the experimental data. We performed additional parameter optimisation on EDS2, with the resultant fits shown in Figure 3.16 (solid black lines). This re-optimisation allows us to recapture

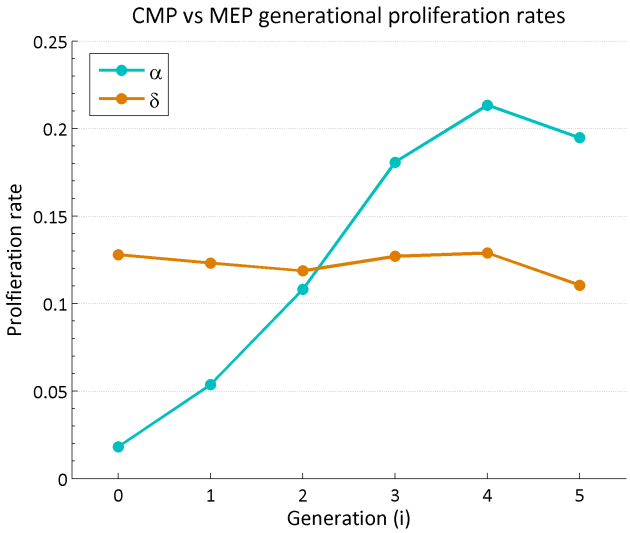


FIGURE 3.14. **Graphical representations of the proliferation rates of CMP and MEP cells.** The fitted proliferation rates for the CMP (blue line) and MEP (orange line) cells in generations 0 through to 5. The proliferation rate for CMP cells appears to increase with each subsequent generation before plateauing, whereas the proliferation rate for MEP cells remains relatively constant with each generation.

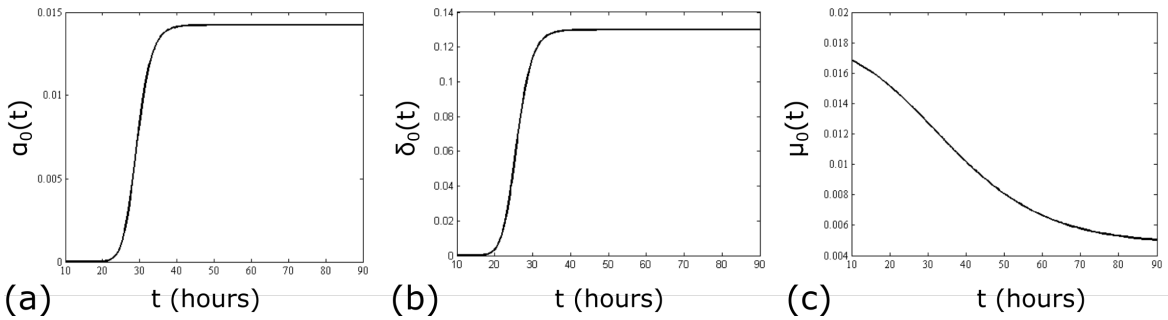
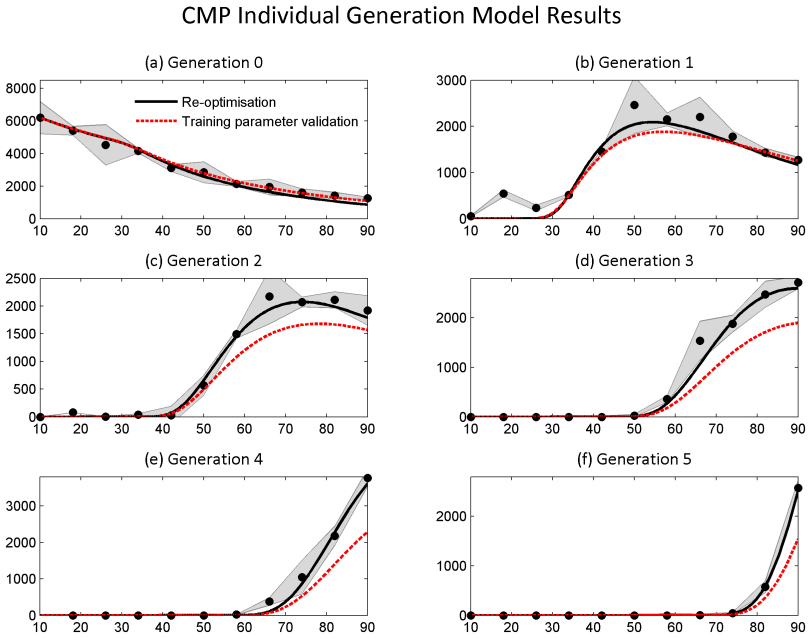
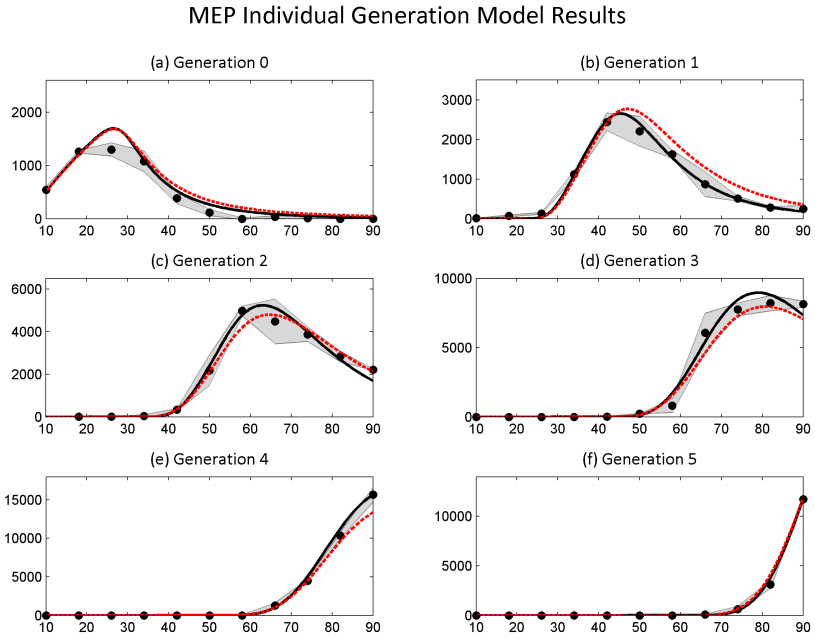


FIGURE 3.15. **Resultant fitted logistic functions for the proliferation rates of G(0) CMP and MEP cells and the maturation rates of G(0) CMP cells to MEP cells.** (a) Fitted proliferation rate of G(0) CMP cells, found by fitting the parameters in Equation 3.5. (b) Fitted proliferation rate of G(0) MEP cells, similarly found by fitting the parameters in Equation 3.6. (c) Fitted maturation rate of G(0) CMP to G(0) MEP cells, found by fitting the parameters in Equation 3.7.



[a]



[b]

FIGURE 3.16. **DDE (adapted Smith-Martin) model parameter fitting validation.** Validation of parameter fitting to the adapted Smith-Martin model, using a parameter set fitted to EDS1 to predict the population dynamics EDS2. (a) CMP counts for the Experimental dataset 2 (black dots show mean of three cell cultures, shaded region shows the standard deviation) with the prediction resulting from the parameters fitted to EDS1 (dashed red line). Included is a re-optimisation of the model to EDS2 (solid black line). (b) as (a) but for MEP cell counts.

TABLE 3.4. **Optimised parameters for the reduced DDE (adapted Smith-Martin) model of cell proliferation and differentiation.**

Parameter	Value
CMP proliferation, P_α	0.018
α_1	0.054
α_2	0.106
α_3	0.178
α_4	0.208
α_5	0.195
Initial maturation, μ_I	0.016
Final maturation, μ_F	0.004
MEP proliferation, P_δ	0.126
δ_1	0.121
δ_2	0.115
δ_3	0.121
δ_4	0.121
δ_5	0.11
CMP initiation delay, τ_c	30.4 hours
MEP initiation delay, τ_m	27.7 hours
CMP proliferation delay, Δ	9.8 hours
MEP proliferation delay, Γ_0	14.5 hours
Γ_1	9.8 hours
Γ_2	8.0 hours
Logistic function parameters, Q_α	0.58
R_α	0.44
Q_δ	0.56
R_δ	0.43
Q_μ	0.618
R_μ	0.065

the population dynamics, with the parameters varying no more than 5% from those fitted to EDS1. This difference in the parameters can be attributed to small differences in the culturing conditions for the population in EDS2, and further validates our use of the adapted Smith-Martin model.

We performed sensitivity analysis on our DDE model, to test the robustness of the parameter identification methods we used, and to quantify the sensitivity of the model to measurement error and also parameter estimation error. Simulations involved systematically adding increasing levels of noise to a baseline dataset (the resultant fit shown in Figure 3.12) and attempting to fit the data, as well as adding increasing levels of noise to our initial 'guesses' for the optimisation routines and attempting to reach the same optimal solution. These *in-silico* experiments were conducted on sets of 1000 'simulated data', generated by adding independent normally distributed

TABLE 3.5. **Results from robustness and system sensitivity analysis of DDE model.** To test the optimisation robustness of the system 7 cases were tested (cases $(a)-(g)$) with increasing simulated data error (DE), which is comparable to the experimental error, and increasing initial parameter guess error (GE) to analyse how a less accurate prediction can still reach the same solution. To test system sensitivity 4 cases (cases $(h)-(k)$) were tested increasing the GE and measuring the data fitting error.

Analysis	Parameter shift (%)	Fitting error (# cells)
a $\pm 3.07\%DE$	0.82	75.83
b $\pm 5\%DE$	0.92	102.04
c $\pm 10\%DE$	2.14	205.03
d $\pm 3.07\%DE \& \pm 5\%GE$	0.79	102.88
e $\pm 3.07\%DE \& \pm 10\%GE$	1.12	114.43
f $\pm 10\%DE \& \pm 5\%GE$	3.17	222.14
g $\pm 10\%DE \& \pm 10\%GE$	2.95	230.95
h $\pm 4\%DE \& \pm 4\%GE$	3.13	99.21
i $\pm 4\%DE \& \pm 8\%GE$	2.74	105.2
j $\pm 4\%DE \& \pm 12\%GE$	3.66	149.7
k $\pm 4\%DE \& \pm 16\%GE$	3.25	167.8

random errors to known solutions of the DDE system and refitting the parameters using the optimisation routines described in Section 3.3.3, providing the means for us to test the system over a range of increasing errors. The results of these experiments are shown in Table 3.5. The first tests consisted of adding error (drawn from a normal distribution, $N(0,\sigma$ where $sigma$ is simlulated data error (DE)) to the known solution so as to make it comparable with the experimental data, using this 'simulated data' to identify how far from the known solution the fitted parameters shift. The addition of 3.07% data error (comparable with the standard error in the experimental datasets) resulted in a mean shift of 0.829% in parameters (Table 3.5, analysis (a)), and an indistinguishable qualitative change in time-courses. Increasing the simulated measurement error to 10% resulted in a shift of only 2.14% from the experimentally-derived parameters (Table 3.5, analysis (c)). Furthermore, we investigated the effect of combining simulated measurement error with initial parameter estimation error (presenting a worst-case scenario, as parameter estimation is, in practice, aided by visual investigation of qualitative properties, not applied in these tests). Errors of 10% measurement and 10% estimation error resulted in a shift in parameters of 2.95% (Table 3.5, analysis (g)). The extensive datasets used for these *in silico* experiments provided us with a clearer understanding of the local parameter space of our DDE system. In the latter worst case scenario, the resulting parameter identification still gave quantitatively and qualitatively accurate fits to the experimental data set (black lines in Figure 3.16).

To garner a more extensive view of the sensitivity of the parameter optimisation to errors in

the initial parameter estimate, we tested an additional set of 600 *in silico* experiments, shifting our initial estimates for each parameter further from the true value and attempting to optimise to experimentally comparable simulated data (Table 3.5, analysis h-k). In this case, increasing the initial parameter estimation error from 4% to 16% resulted in a quantitative change in data fit from 99 cells to 168 cells, and an increase in 0.68% in fitting error.

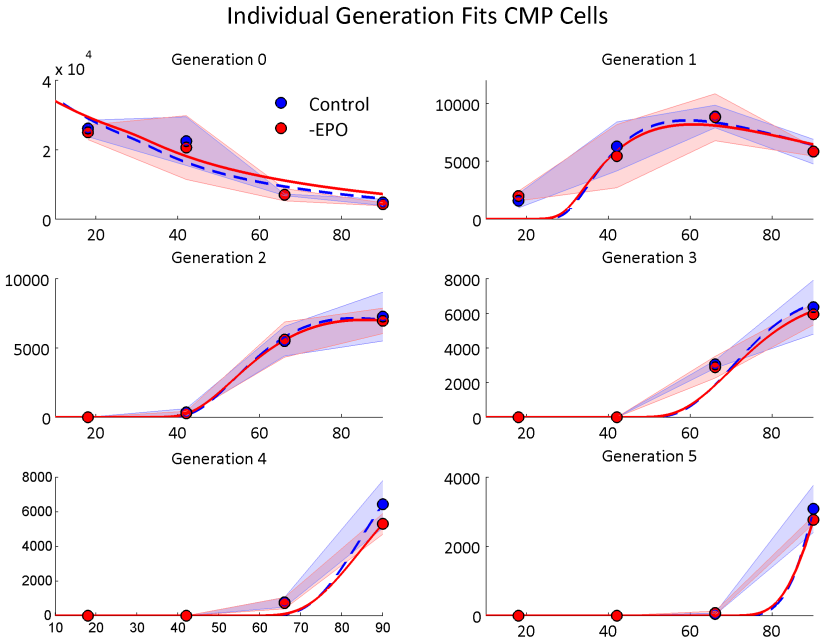
We thus conclude that our parameter optimisation methods are robust to noise, both in experimental measurement and initial parameter estimates.

3.6 Elucidating the Role of Erythropoietin (epo)

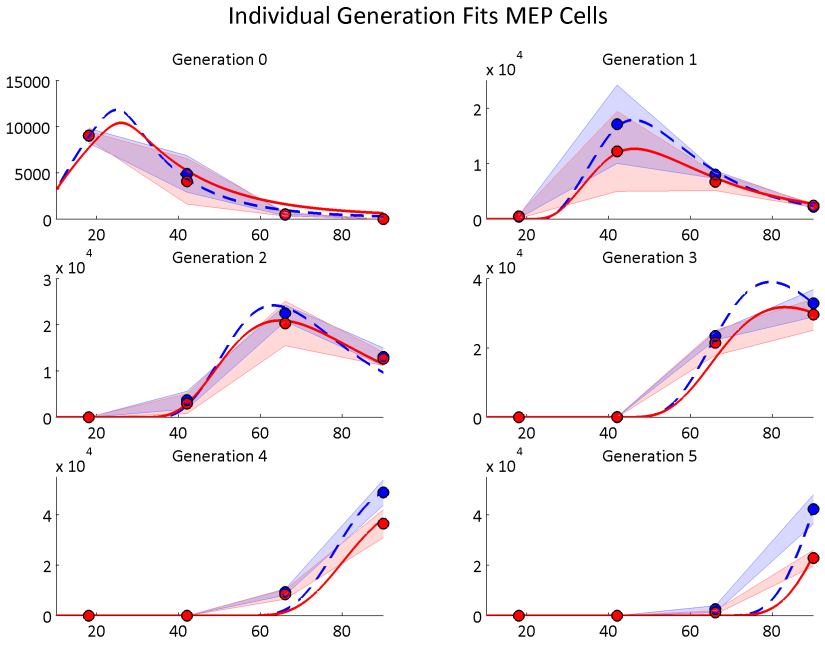
To investigate a possible role of epo on the proliferation and maturation of cells *in vitro* we fitted the adapted Smith-Martin model to two additional culture experiment datasets, a control EDS3 (+epo) and EDS4 (−epo) dataset. In the earliest stages of erythropoiesis in vitro, CD34+ cells were isolated and cultured with/without epo: Figure 3.17 shows the generational counts for CMP and MEP cell cultures in the presence (blue dots) and absence (red rods) of epo. After 1 day in culture, prior to the onset of proliferation, 20% CMP have matured to MEP under both conditions. This suggests that the CMP to MEP transition, during the first 24 hours, is not responsive to epo in G(0). Over 9 days in culture, there is a dramatic difference in the number of MEP in control and minus epo cultures and by day 9 there is over 20 times more MEP produced in cultures with epo.

Long-term MEP reduction in the absence of epo could be due to combined maturation or proliferation defects, or both. To resolve these possible epo effects, we fitted our model on such datasets; results are shown in Figure 3.17 (generational data and corresponding model fitting) and Figure 3.19 (total cell counts in culture). Parameter fitting shows that the proliferation rate of MEP (δ) is on average 21% lower in the absence of epo, whereas there a lesser change (approximately 9%) in the proliferation rate of CMP (α), as shown in Table 3.6. The traces from CFSE tracking also support this result in that a distinct delay in generation progression for MEP is detectable in cultures without epo, although differences are below the level of detection at earlier time points (not shown). Also, CFSE tracking shows no significant effect of epo on CMP generation occupancy; this result, taken together with the unchanged proliferation rates in CMP, indicates that CMP proliferation is not epo sensitive. The maturation rate for G(0) is not significantly different with or without epo, and the final maturation rate is even higher in absence of epo, suggesting that epo effects are not mediated by altered maturation.

To further demonstrate that maturation of CMP to MEP is insensitive to epo, we have used CFSE tracking to calculate the number of CMP that do not mature, taking into account the number of cell divisions, to look at how the maturation of CMP in the initial G(0) population drops off as time progresses. We noticed that, as the experiment progresses, the maturation of G(0) CMP reduces, as shown in Figure 3.18 (dots represent experimental data). The results from the model



[a]



[b]

FIGURE 3.17. **Generational population counts and model fitting for control and −epo datasets.** Generational data from culture experiments with/without epo and the corresponding model fittings. (a) Generational counts of CMP for the two culture cases: a control (+epo) case represented by blue dots (mean of three experiments) and blue shaded region (standard deviation), and a culture without epo (−epo) represented by red dots (mean of three experiments) and red shaded region (standard deviation). (b) As in (a) but for MEP cell population.

TABLE 3.6. **Optimised parameters for the reduced DDE (adapted Smith-Martin) model for +/- epo cultures.** Table of model parameters fitted to experimental data of HHSC cultures in the presence of epo (EDS3) and in the absence of epo (EDS4). Difference between fitted parameters are shown as a %Δ.

Parameter	+epo (EDS3)	-epo (EDS4)	%Δ
CMP proliferation, α_0	0.014	0.012	-14.3
α_1	0.045	0.042	-6.67
α_2	0.098	0.086	-12.2
α_3	0.198	0.174	-12.1
α_4	0.264	0.255	-3.41
α_5	0.299	0.269	-10.03
Initial maturation, μ_I	0.025	0.020	-20.0
Final maturation, μ_F	0.003	0.005	66.7
MEP proliferation, δ_0	0.118	0.088	-25.4
δ_1	0.121	0.114	-5.79
δ_2	0.158	0.119	-24.7
δ_3	0.181	0.146	-19.3
δ_4	0.279	0.206	-26.2
δ_5	0.269	0.199	-26.02
CMP initiation delay, τ_c	29.7 hours	28.8 hours	-2.9
MEP initiation delay, τ_m	26.3 hours	25.8 hours	-1.9
CMP proliferation delay, Δ	10.7 hours	10.6 hours	-1.2
MEP proliferation delay, Γ_0	14.6 hours	12.5 hours	-14.69
Γ_1	10.5 hours	9.5 hours	-9.67
Γ_2	10.3 hours	10.9 hours	5.59
Logistic function parameters, Q_α	0.58		
R_α	0.44		
Q_δ	0.56		
R_δ	0.43		
Q_μ	0.618		
R_μ	0.065		

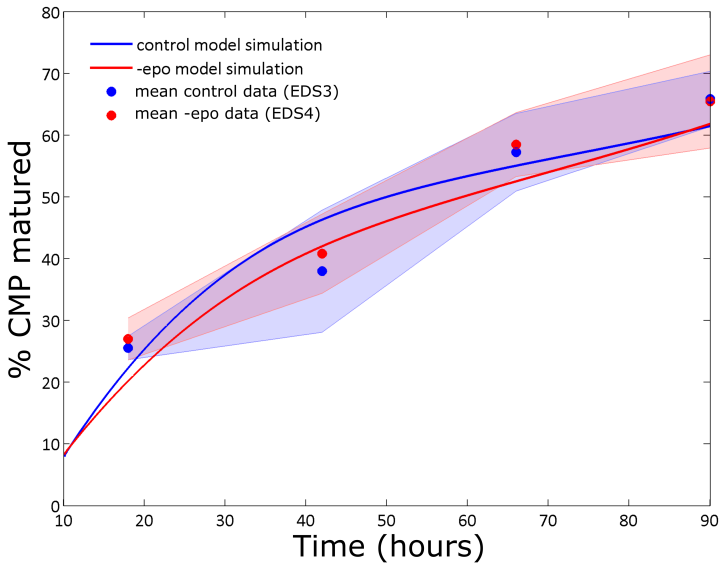


FIGURE 3.18. Maturation dynamics of CMP cells in generation 0 (G(0)), showing that the percentage of CMP cells in G(0) that have matured to MEP begins to plateau. The blue/red dots are the percentage of CMP cultured with/without epo that has matured, respectively. The blue/red lines show the prediction of our model fitting for CMP maturation in G(0) for +/- epo cultures, respectively.

fitting indicate that maturation after G(0) has an insignificant effect on population dynamics, as shown by the parameters fitted to both the culture with/without epo in Table 3.6. Taken together with the finding above that MEP formation in G(0) is indistinguishable between control and minus epo cultures, and results from the mathematical model, we have strong evidence that neither the rate nor the bias of the CMP to MEP transition is influenced by epo.

Figure 3.18 shows how the proportion of CMP cells from the initial population that have matured begins to plateau as we reach 90 hours of culture. The difference between the maturation percentage of the two populations, with and without epo, is insignificant in comparison to the standard deviation of the data. This suggests that the role of epo is minimal in this early stage of erythroid cell culture. Our model does fit this trend towards a plateau as the experiment goes on, in part helped by the time-dependent maturation of CMP cells.

We finally used the calculated parameters for 4-day cultures with and without epo to simulate MEP population expansion to day 8 (Figure 3.20). In order to extend the model to cover the extension to day 8, we had to include a further 6 generations of CMP and MEP.

To simulate the extrapolation model, we chose a linear relationship between the proliferation rates shown in Table 3.6; the chosen parameter values for MEP proliferation rates are shown in Table 3.7. For +epo (control) cultures, there is a good agreement between the predicted cell numbers and experimental data; for cultures without epo, there is good agreement for the first 5-6

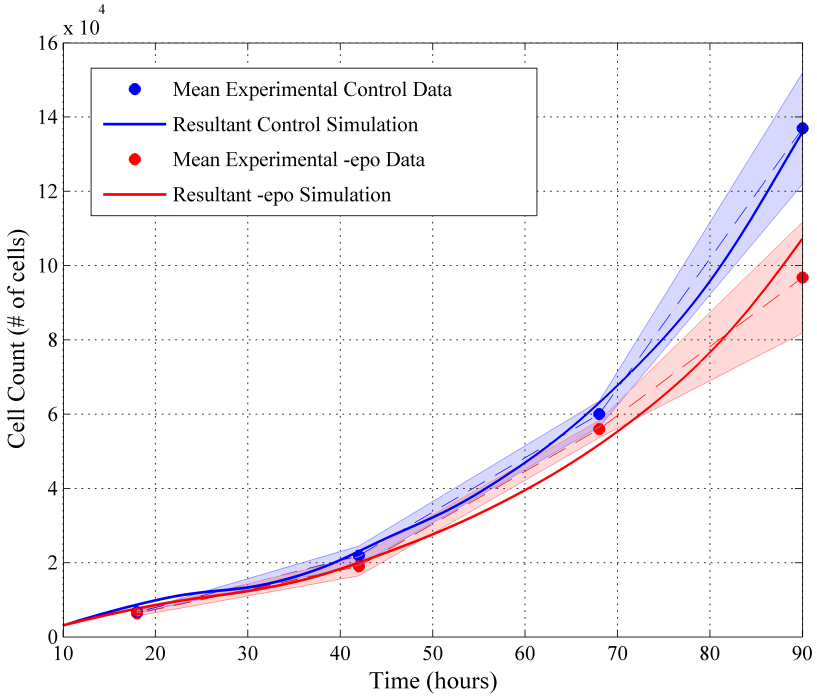


FIGURE 3.19. **Total cell-count comparison results with/without epo.** Model fitting simulations for the total cell counts of control (blue) and −epo (red) cultures compared to experimental data (dots). Shaded regions represent standard error in the experimental data

days (up to time 130 hours in Figure 3.20), and then the experimental data show a significantly slower proliferation than predicted.

We found a marked change in the population dynamics after day 5 (Figure 3.20) when simulating MEP amplification to day 8 in the absence of epo using the proliferation rate derived from the first 4 days of culture. This suggests that there are two MEP stages with different epo sensitivity: an early stage (day 1 to day 4) where proliferation is lesser affected by epo, and a later stage (day 4 to day 8), which is increasingly dependent on epo for proliferation. The increased resolution of cell behaviour parameters given by the mathematical model provides a valuable tool to further investigate both the mechanism of lineage restriction and proliferation behaviour of haematopoietic progenitors in culture.

3.7 Discussion

In this chapter, we have developed a mathematical model to allow us to capture the differential effects on proliferation and differentiation of the culture hormone, erythropoietin (epo), on human haematopoietic stem cells (HHSCs). We aimed to determine whether its inclusion throughout

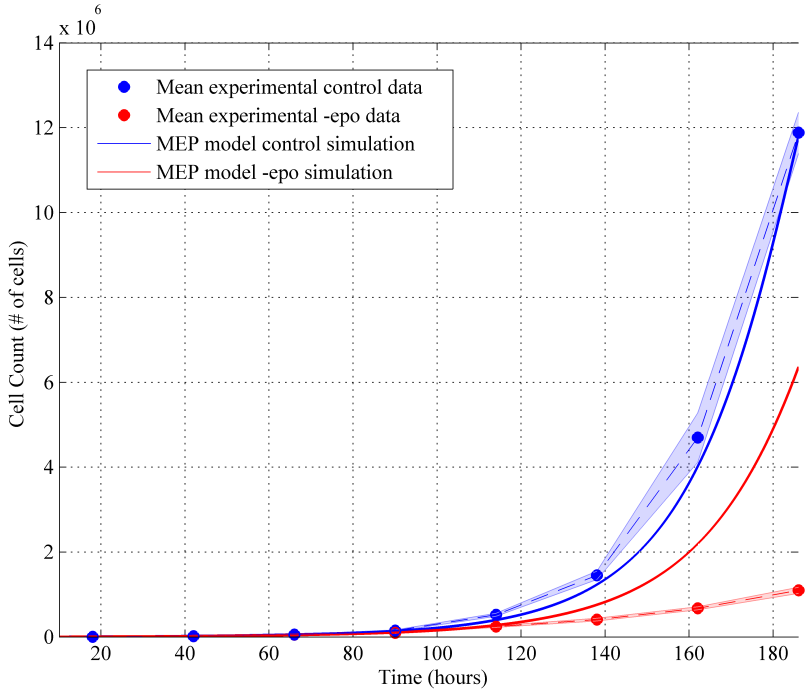


FIGURE 3.20. Extrapolation of MEP cell counts to Day 8 (186 hours) using best-fit parameters in Table 3.4 for both control and −epo populations. Additional proliferation rates ($\delta_{(6-13)}$) predicted by assuming a linear relationship between the optimised proliferation rates, found from the Day 0−4 fitting for G(0) to G(5) MEP, allowing us to extrapolate forward for G(6) onwards.

culture is necessary or whether it can be introduced at a later stage where it is well documented to be required [102]. By adapting the Smith-Martin model [155], a formalism which uses delay differential equations (DDEs) to describe cell-cycle progression and thus cell proliferation, we were able to simultaneously describe the differentiation of HHSCs by incorporating the differentiation step between two instances of the Smith-Martin model. Using such an extended DDE model for cell proliferation and differentiation, we were able to effectively capture the growth dynamics of HHSC cultures, as shown in Figure 3.12, with a root mean square error (RMSE) of approximately 5%. We were able to show, as detailed in the parameters in Table 3.4, that the proliferation of less mature CMP cells increases with each generation, whilst the more mature MEP cells maintain a relatively constant proliferation rate. This was unexpected, as it suggests that the formation of a maximal MEP colony has a greater dependency on the proliferation of the less mature cell type and not their own cell type. Similarly unexpected was the results that the maturation of CMP into MEP cells is weighted towards the early stage of culture, with the majority of maturation occurring before we reach the midpoint of the culture experiment, with maturation reducing by 75% between the start and the end of the culture experiment (i.e. after 96 hours in culture).

TABLE 3.7. **Proliferation rates for MEP in DDE model extrapolation experiments.** Table of parameters chosen for the extrapolation validation of our adapted Smith-Martin model. Parameter values are linearly increasing with the relationship determined between the parameters shown in Table 3.6.

Parameter	+epo (EDS5)	-epo (EDS6)
MEP proliferation, δ_6	0.314	0.234
δ_7	0.350	0.259
δ_8	0.386	0.284
δ_9	0.422	0.309
δ_{10}	0.458	0.334
δ_{11}	0.494	0.359
δ_{12}	0.530	0.384
δ_{13}	0.566	0.409

Following the fitting of the adapted Smith-Martin model to experimental data, we used the same framework to analyse the effects of epo on the cultured population. Although it is well established that epo is required for proliferation of late erythroid progenitors, a possible influence of epo on earlier progenitors had not been previously described. To test this influence, we determined the proliferation and maturation rates using our model for both a population in the presence of epo and a population cultured in the absence of epo. Figures 3.17 and 3.19 show the population counts in both conditions. We showed that, from the initial stages of megakaryocyte/erythroid lineage commitment, cells are epo responsive and proliferation is slower in the absence of epo, although it appears that CMP cells are lesser affected by epo than MEP cells. The model allowed a high-resolution quantification of proliferation rate that revealed the sensitivity of MEP to epo, by disentangling the independent contributions of proliferation and maturation to population dynamics. This means of quantification can be highly beneficial to designing optimal protocols for amplifying erythroid cells in culture and producing red cells for clinical transfusion.

As a final verification for our model, we used the model parameters derived from the first 4 days of culture (Table 3.4) to extrapolate dynamics over 8 days of *in-silico* cell culture. We showed that the derived proliferation rates faithfully reproduce the experimental data (Figure 3.20). This confirms the power of the model to describe erythroid culture and to predict changes to the dynamics. However, when we simulated the cell culture to day 8 in the absence of epo using the proliferation rates derived from the first 4 days, we observed a marked change after day 5, depicted by the red line in Figure 3.20. This suggests that there are two stages in MEP culture each with different epo sensitivity: an early stage (days 1–4) with a small increase in proliferation rate in response to epo, and a later stage (days 4–8), which is increasingly dependent on epo for proliferation.

The ability to disentangle cell behaviour parameters, namely related to proliferation and maturation, given by the mathematical model, provides a valuable tool to further investigate both the mechanism of lineage restriction and proliferative behaviour of haematopoietic progenitors in culture. In this chapter, we have not considered the stochastic nature of cell culturing and the effects this may have on emergent dynamics. In the following Chapter we will look at the culturing of mouse embryonic stem cells (mESCs), using stochastic differential equations to describe the bistability of key genes linked to proliferation and differentiation, and incorporate this additional stochasticity into an agent-based modelling framework to capture emergent behaviours.

Chapter 4

Agent-based Modelling of Mouse Embryonic Stem Cell Signalling Pathway and Proliferation Dynamics

Embryonic stem cells' ability to maintain a pluripotent (undifferentiated) state during development is essential to the healthy maintenance of living organisms. These cells play a key role in the repair and replacement of organ and epithelial tissues, due to their capacity to differentiate into a multitude of different cell types. However, the process by which stem cells, and here specifically mouse embryonic stem cells (mESCs) regulate their pluripotency (differentiative capacity) in culture is not fully understood. This is especially true with regards to the effects of culture medium on cell pluripotency. It is possible to achieve varying levels of pluripotency in an isogenic population of cells if they are cultured in different media; however, the gene expression mechanisms by which the state of cells, and the dynamics underlying cell fate, is affected require additional investigation.

This chapter extends the modelling techniques used in Chapter 3, describing the development and chemical dependencies of human haematopoietic stem cell culture population dynamics through the use of delay differential equations, by introducing an agent-based approach to model heterogeneous stem cell populations. Where the modelling in Chapter 3 used global population properties to describe cell dynamics, here we look at multiple biological and mathematical scales. We considered the internal dynamics of relevant signalling pathways in mESCs, as well as their links to the cell cycle, and emergent pluripotency phenotypes. The use of agent-based modelling

provides an understanding of the system on a population level, allowing us to investigate any dynamics previously hidden in single cell simulations. We investigated an extended gene regulatory network (GRN) [68] which has been shown to govern the pluripotent state of cultured mESCs. Using agent-based modelling we elucidated the genetic mechanisms by which this GRN is linked to the mESCs cell cycle, and thereby we were able to show that cell proliferation is linked to the levels of MycN in cells, providing a novel mechanism to explain the effect of culture conditions on mESC pluripotency.

The work shown in this Chapter has been published as an article, titled "An extended model for culture-dependent heterogenous gene expression and proliferation dynamics in mouse embryonic stem cells" [68]. The derivation of the extended network in Section 4.2 was conducted by Simon Godwin (current PhD student in Engineering Mathematics, University of Bristol), the culturing experiments were conducted by Dr Elisa Pedone, and the paper was written in collaboration with Dr Martin Homer, Dr Alexander G. Fletcher and Dr Lucia Marucci. I developed the agent-based model, tested the response of the system to changes in Nanog and MycN concentrations and conducted *in silico* cell-sorting experiments.

Section 4.1 provides relevant biological background, providing context for the derivation of a new GRN and further investigation through agent-based modelling. Section 4.2 details the derivation and development of the extended GRN for the culture of mESCs. Section 4.3 explains the development of the agent-based model of mESC culturing. Section 4.4 shows the results of the agent-based models developed, including the fitting of the models to experimental data. Section 4.5 discusses the key results from our modelling work and possible extensions that could be considered for further analysis in the area of agent-based modelling of mESC dynamics.

4.1 Background: mESC Pluripotency and Culture Conditions

In this chapter, we consider mouse embryonic stem cells (mESCs) and the differential effects of culture conditions on their pluripotency. Understanding how to control the pluripotency/proliferative capacity of stem cells is key to formulate improved culture protocols. mESCs are isolated from the mouse inner cell mass and have the capacity to be pushed into specific differentiation paths by different stimuli *in vitro*, providing the possibility of made-to-order cell populations of any cell types, all from a single starting cell population.

mESCs can be cultured *in vitro*, typically in a culture medium named serum/LIF [23]. This medium contains serum response factors, transcription factors which directly regulates genes responsible for stem cell fate regulation, and cytokine leukaemia inhibitor factors (LIF). The combination of these factors affects cell growth due to the inhibition of cellular differentiation, resulting in a slowly proliferating but wholly pluripotent cell population. It has been reported [81] that this medium creates a state of heterogeneous expression and temporal fluctuations of pluripotency factors and regulators which include Nanog, Rex1, Stella, Esrrb and β-catenin. It

has been suggested that many of the metastable mESC genes are regulated by Nanog, a master regulator of pluripotency and development [23]. The relative levels of Nanog expressed by the cells affect their propensity for differentiation, with cells exhibiting low levels of Nanog more likely to differentiate than those exhibiting high levels of Nanog. Cultured mESCs are able to switch between these Nanog states, with the fluctuation likelihood linked to the culture medium used; thus, if we want to maintain a fully pluripotent population, we would like to keep the cells in a homogeneous Nanog high state (i.e. ground state of pluripotency). To this end, an alternative culture medium has been developed [183] which is capable of reducing the heterogeneity of pluripotency gene expression in cultured mESCs. This medium, named 2i/LIF, is serum free and instead incorporates two chemical inhibitors, PD (PD0325901, MEK inhibitor) and Chiron (CHIR99021, glycogen synthase kinase-3 (Gsk3) inhibitor). PD is a protein kinase inhibitor which selectively binds to and inhibits MAPK/Erk phosphorylation and effectively reducing the inhibition of Nanog by Erk. Chiron functions as an activator of the Wnt signalling pathway by promoting β-catenin within the cell, as well as reducing the TCF3 mediated inhibition of Nanog.

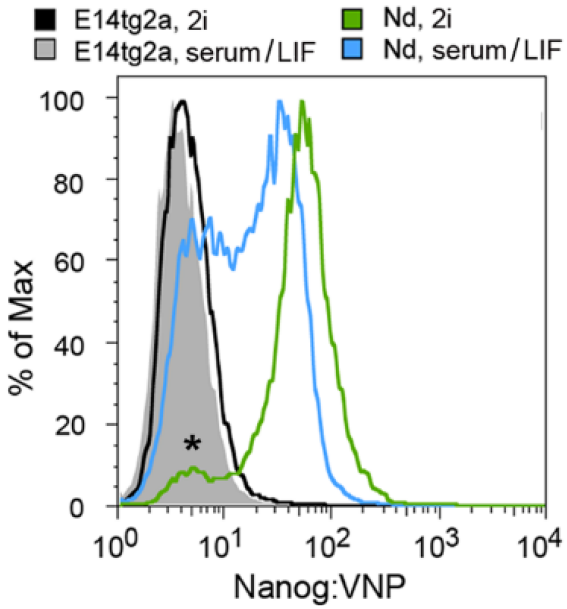


FIGURE 4.1. **mESC Experimental data showing the differential Nanog state composition in Serum/LIF vs 2i/LIF.** Histogram showing the Nanog expression, measured by a fluorescent reporter (VNP) in the Nd cell line (a novel ES cell line containing a transgenic bacterial artificial chromosome (BAC))[3] grown in serum/LIF (blue) and in 2i/LIF (green). The grey region and black line represent the negative control cells (E14tg2a) which are grown in the same conditions. In serum/LIF, roughly 55% of the cells express Nanog, whereas, when cultured in 2i/LIF, this value increases to around 90%. Despite this change in expression levels, a population of mESCs with no or low levels of Nanog is always observed (*). Figure reproduced from Abranches *et al* [3].

Culturing mESCs in serum/LIF results in a heterogeneous population expressing low levels of Nanog almost as readily as high levels of Nanog (55% Nanog-high(NH) and 45% Nanog-low (NL) [23]), and so almost half of the cells within the population have a propensity for differentiation. In order to reduce the prevalence of NL cells, we can compare with the expression found in mESCs cultured in 2i/LIF medium. Although initially reported as a medium that would completely eliminate the NL sub-population seen in serum/LIF culture [23], more recent data [2, 3] suggests that there remains a small sub-population of NL cells that are not in the ground state pluripotency (90% NH and 10% NL [183]). Figure 4.1 shows a comparison of Nanog expression levels in mESCs cultured in both the serum/LIF (blue histogram) and the 2i/LIF media (green histogram) [3]. In order to understand why the NL sub-population remains within the cell population cultured in 2i/LIF, we developed an extended model of the gene regulatory network (GRN), that includes genes key to maintaining pluripotency and restricting proliferation, which are also affected by the two inhibitors present in 2i/LIF [68].

4.2 Mathematical Modelling of the mESC Population

The model we developed extends the GRN model described in Herberg *et al* [84], which captures mESC dynamics in serum/LIF. The modelled GRN is made up of the core pluripotency transcription factors Oct4, Sox2, FGF4-Erk, Nanog and Rex1, and can recapitulate Nanog bistability in serum/LIF. This basic network can be adapted to incorporate the first of the two inhibitors, PD, due to its effect on FGF4-Erk inhibition of Nanog, however, it does not take into account the second inhibitor Chiron present in 2i/LIF medium.

In order to identify possible genes for inclusion in an extended network, an investigatory workflow was applied, reproduced in Figure 4.2 [68]. This workflow involved re-analysing published RNA-sequencing data of mESCs cultured in serum/LIF and 2i/LIF [114], and performing gene ontology (GO) analysis, using the Database for Annotation, Visualization and Integrated Discovery (DAVID) platform [41], on the genes that were found to be differentially expressed in the two media. The genes were then categorised by their associated biological processes (BPs), with those relevant to development and differentiation, proliferation, cell cycle, morphology or cell death retained. Those resulting genes were then filtered to include only those which interact with the initial core network pluripotency factors (i.e Oct4, Sox2 and Nanog) [83]. This filtering was done using CODEX [145], a compendium for mESC transcription factors, with genes regulating at least one of the core factors considered. From this analysis, MycN, Rest and Prdm14 were identified as key inclusions in the extended network, with these factors also reported to significantly affect mESC pluripotency [68]. To add the directions of gene interaction (i.e activation or repression), the NIA mESC cell bank was used[128]; this data-set was generated by measuring the global expression of mESCs following perturbation of various transcription factors. The core network interactions were kept the same as in Abranches *et al* [3]. The NIA bank

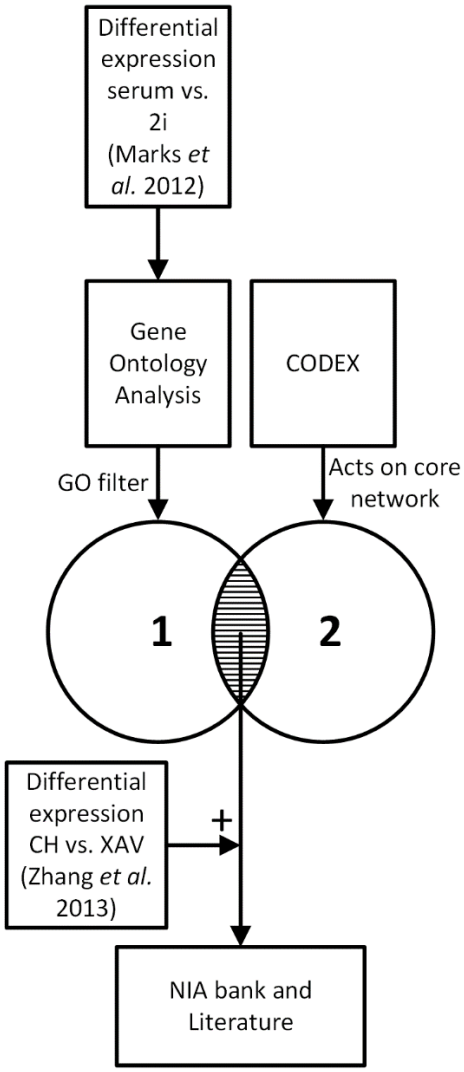


FIGURE 4.2. **Extended Nanog network gene analysis workflow.** Diagram representing the workflow applied to refine the base Nanog regulatory network in Abranches *et al* (2014) and to identify additional genes crucial for pluripotency maintenance and transcription heterogeneity for inclusion in the extended model. Reproduced from Godwin *et al*, 2017 [68]

analysis found that the Oct4-Sox2 heterodimer activates Nanog, whilst it also simultaneously, indirectly, represses Nanog through the FGF4-Erk module. There was no data in the NIA bank for the additions of Rest and Prdm14, and so only interactions reported in the literature were included for the interactions of those genes.

The resulting extended network is shown in Figure 4.3: it included the core GRN (Oct4-Sox2 heterodimer, FGF4-Erk and Nanog) as well as the newly identified genes MycN, Rest and Prdm14 and the interactions between them. The latter genes were added as they are differentially expressed in 2i/LIF vs Serum/LIF whilst also forming feedback loops with the core network. As

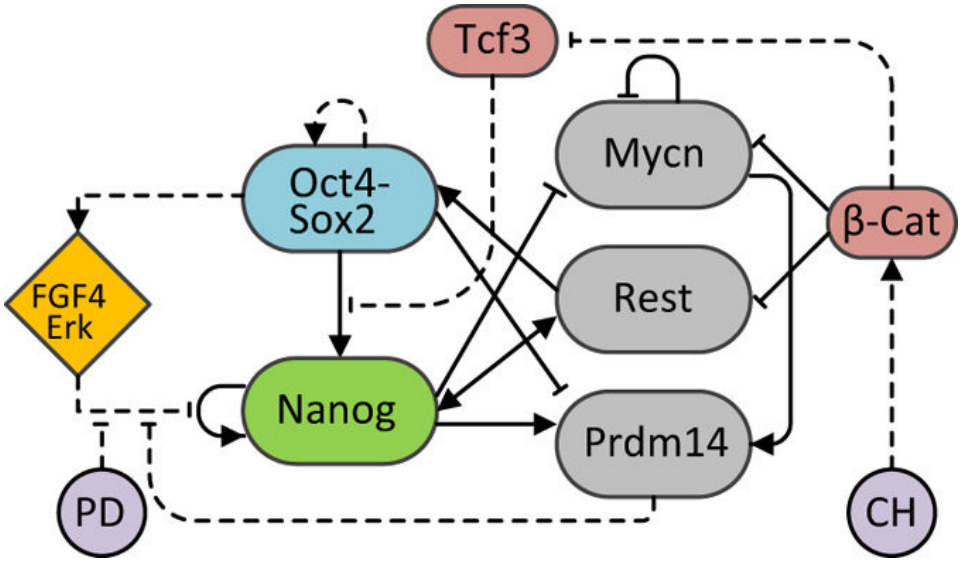


FIGURE 4.3. **Network diagram for mESC pluripotency GRN.** Diagram showing the extended GRN to capitulate the pluripotency dynamics expected in mESCs in both serum/LIF and 2i/LIF cultures. The solid lines indicate transcriptional interactions and the dashed lines indicate non-transcriptional interactions. Reproduced from Godwin *et al* (2017).

in Figure 4.3, there is an indirect activation of Nanog through the MycN-dependent activation of Prdm14, which in turn prevents FGF4-Erk based repression of Nanog. The effects of Chiron within the culture medium is incorporated via the inclusion of β-catenin and the transcription factor Tcf3, which in turn represses Oct4-Sox2 activation of Nanog.

The network depicted in Figure 4.3 can be described in terms of stochastic differential equations (SDEs) (Equations 4.1 to 4.9). They describe the temporal dynamics of the network genes (N=Nanog, OS=Oct4-Sox2, M=MycN, PR=Prdm14, R=Rest, E=Fgf4/Erk, B=β-catenin, $Tcf3$=Tcf3), specifically they are able to replicate the Nanog state-switching dynamics present in cultured mESCs due to the incorporation of the noise term (ϕ). The parameters for network SDEs are presented in Table 4.1, with the parameters estimated by performing 50 simulations, each for 1000 individual cells as independent solutions of the SDE model, in order to consistently describe the flow cytometry data of Nd mESCs cultured in both serum/LIF and 2i/LIF [3].

$$p = \frac{s_{p1} * k_{p1}}{k_{p1} + [PR]} + \frac{s_{p2} * k_{p2}}{k_{p2} + [PD]}, \tag{4.1}$$

$$\frac{d[OS]}{dt} = n * \left(\frac{s_{1,2} * [OS]^2}{(k + [OS])^2 * dO * dS} + \frac{s_3 * [R]}{k + [R]} - dOS * [OS] \right) + \phi(0, \sigma_{os}) * [OS], \tag{4.2}$$

$$\frac{d[N]}{dt} = n * \left(\frac{s_4 * N^2}{k^2 + [N]^2 + p * [E]} + \frac{s_5 * [OS]}{k + [OS] + [Tcf3]} + \frac{s_6 * [R]}{k + [R]} - dN * [N] \right) + \phi(0, \sigma_N) * [N], \tag{4.3}$$

$$\frac{d[M]}{dt} = n * \left(\frac{s_7 * k}{k + [M]} + \frac{s_8 * k}{k + [N]} + \frac{s_9 * k_m}{k_m + [B]} - dM * [M] \right) + \phi(0, \sigma_M) * [M], \tag{4.4}$$

$$\frac{d[PR]}{dt} = n * \left(\frac{s_{10} * k}{k + [OS]} + \frac{s_{11} * [N]}{k + [N]} + \frac{s_{12} * [M]}{k + [M]} - dPR * [PR] \right) + \phi(0, \sigma_{PR}) * [PR], \tag{4.5}$$

$$\frac{d[R]}{dt} = n * \left(\frac{s_{13} * [N]}{k + [N]} + \frac{s_{14} * k_r}{k_r + [B]} - dR * [R] \right) + \phi(0, \sigma_R) * [R], \tag{4.6}$$

$$\frac{d[E]}{dt} = n * \left(\frac{s_{15} * [OS]}{k + [OS]} - dE * [E] \right), \tag{4.7}$$

$$\frac{d[B]}{dt} = n * \left(\alpha_1 - dB * \left(\frac{s_b * k_b}{k_b + [CH]} \right) * [B] \right) + \phi(0, \sigma_B) * [B], \tag{4.8}$$

$$\frac{d[Tcf3]}{dt} = n * \left(\alpha_2 - dTc * \left(1 + \frac{s_t * [B]}{k_t + B} \right) * [Tcf3] \right) + \phi(0, \sigma_T) * [Tcf3]. \tag{4.9}$$

Figure 4.4 shows the resultant distributions of Nanog in simulated cultures of mESCs using the extended network, with the distributions estimated from the 1000 individual solutions to the SDEs. Figure 4.4(a) shows that the model is capable of capturing the serum/LIF distribution of Nanog, with approximately 55% of cells in a NH state, which has been seen in Herberg *et al* [84]. It is also able to capture the distribution of Nanog expected in a 2i/LIF culture as shown in more recent studies [3]. This distribution maintains a much-reduced subpopulation of NL cells (10%) as expected. Prescribing whether a cell is in a NH or NL state switches at the midpoint of the two Nanog peaks (the NL and NH peaks in Figure 4.4).

Although we showed that the SDEs model of the extended GRN is capable of capturing the differential distributions of Nanog seen between the two culture media, to assess the role of cell division, and its coupling to single cell dynamics, we extended the SDE model in an agent-based context. By doing so, emergent dynamics in mESC populations can be captured.

TABLE 4.1. Parameters for the extended Nanog network (equations 4.1-4.9), derived from the network diagram depicted in Figure 4.3 and fitted to experimental data [3]. * molecules/minute; # 1/minute. Parameters reproduced from [68].

Extended Nanog network parameters					
s_{p1}	30*	s_{14}	20*	dM	1#
s_{p2}	30*	s_{15}	2*	dPR	1#
k_{p1}	5	s_b	50*	dR	1#
k_{p2}	15	s_t	20*	dE	1#
$s_{1,2}$	75*	α_1	1*	dB	1#
s_3	1*	α_2	40*	dTc	1#
s_4	37*	k	10	σ_{OS}	0.02
s_5	2*	k_r	0.05	σ_N	0.05
s_6	0.01*	k_m	0.05	σ_M	0.02
s_7	10*	k_t	10	σ_{PR}	0.02
s_8	100*	k_b	1	σ_R	0.02
s_9	20*	dO	1#	σ_B	0.02
s_{10}	15*	dS	1#	σ_T	0.02
s_{11}	30*	dOS	1#	s_{nb}	10*
s_{12}	10*	dN	1#	n	0.005
s_{13}	0.1*				

4.3 Agent-Based Modelling of mESC Signalling Pathway and Proliferation Dynamics

4.3.1 Set-up of the Agent-Based Simulations

Agent-based models provide a means to investigate population dynamics that can arise when you scale up from a single cell to a large population of cells, reducing the homogeneity of the population. Considering each individual cell as an agent, we can include the internal state for each cell depending on the extended GRN for Nanog bistability (shown in Figure 4.3) and use it to elucidate the mechanism coupling the GRN to the cell cycle of mESCs. The focus of our work applying agent-based modelling to this system is to determine the effect of gene-dependent cell-cycle progression on the global population dynamics.

In forming our agent-based model, we consider what additional information we need about the cell culture conditions, as well as the cells proliferation dynamics, to create a model that is well-informed enough to reproduce the key dynamics that we are studying without becoming convoluted. When developing our agent-based model, we considered three basic interconnected modules, corresponding to the different physical scales (levels) involved.

1. 'Subcellular level' network models;

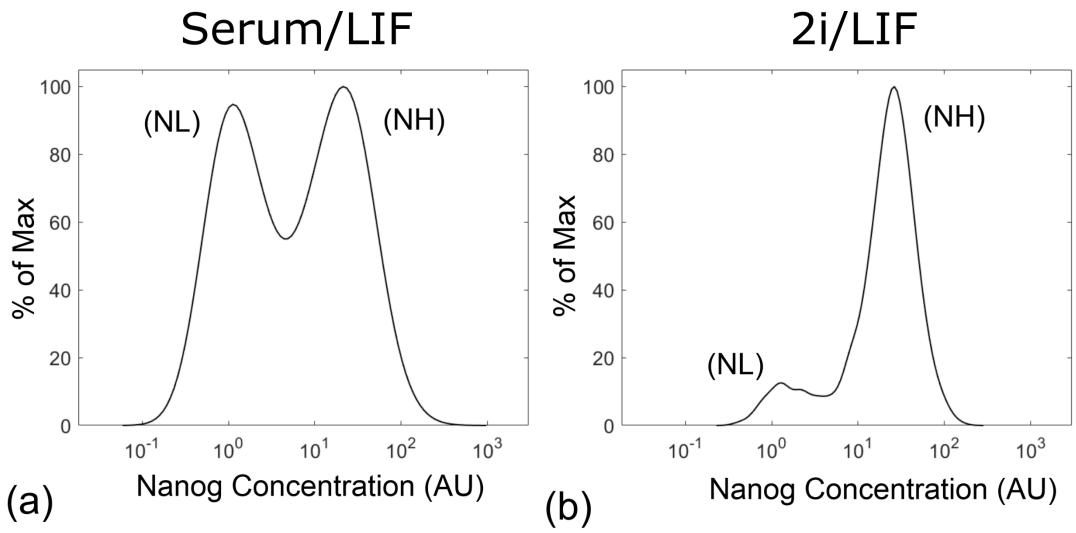


FIGURE 4.4. **Results of the extended network SDE model, showing the differential expression of Nanog, dependent on the culture medium used (serum/LIF and 2i/LIF).** (a) Resultant distribution of Nanog within a simulated population cultured in serum/LIF showing the expected distribution of Nanog (55%NH/45%NL). (b) Resultant distribution of Nanog within a simulated population of mESCs cultured in 2i/LIF, with the expected distribution of Nanog (90%NH/10%NL). GRN was integrated using the Euler-Maruyama method [85]. Reproduced from Godwin *et al* [68].

2. 'Cellular level' cell-cycle models;

3. 'Population level' mechanical models.

These three levels of modelling are all interconnected, with the subcellular network models influencing the behaviour of the cellular level cell-cycle models, which in turn affect the population mechanics. Combining the models in this way creates a multiscale model, capturing the dynamics of the system from a micro- to macro-scale. We chose what models for each of these levels to consider, based on the importance of the level of the overall behaviours we wish to capture.

The agent-based model was developed within the Chaste simulation framework (detailed in Chapter 2 Section 2.2), which allowed for the development of an agent-based model in the modular fashion previously described. The interactions of the modules within the Chaste simulation framework is shown in Figure 4.5, formed of modules corresponding to the three scales described above.

In the case of mESC culture, the most important aspect we consider is the cell-cycle model. It has been recently shown that the length of the cell cycle in mESCs is dependent on the internal state of the subcellular network model, specifically the levels of Nanog exhibited by the cells [164], the combined effect of a Nanog-dependent cell cycle, coupled to proliferation, can affect the overall distributions of Nanog within the population seen in Figure 4.4 [84]. This could be attributed to

a positive correlation between Nanog concentration in individual cells and proliferation, with more rapidly proliferating cells contributing most to the make-up of the cell culture. For our modelling, the mechanics involved in cell-cell interactions are not as important; this is because we can assume that the cell culture has homogeneous growth conditions preventing the occurrence of contact inhibition or increased cell apoptosis (biological processes linked to the mechanistic properties of cells and their interaction with neighbouring cells). This is backed up by the experimental process, with the cells reaching confluence within their growth environment beyond the 4 days we are interested in [68]. The subcellular model provides a framework for the GRN describing the dynamics of gene concentrations, and/or other chemical changes within individual cells. This GRN SDE model can be coupled to the cell-cycle model, the mechanical model, or both; as mentioned, here our GRN only affects the cell-cycle. Using an agent-based model allows us to also fit our parameters to population growth dynamics, with the differential effects of NH/NL cell-cycle length difference on the growth of the population.

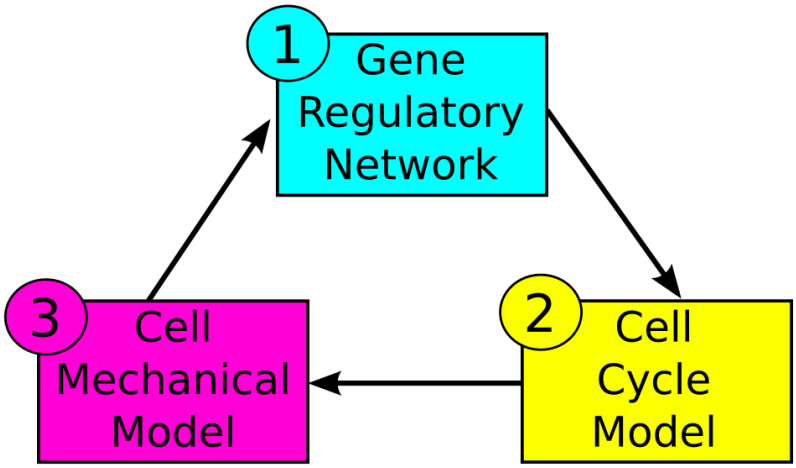


FIGURE 4.5. **Diagram illustrating the modular nature of the Chaste agent-based simulation programming package.** Chaste incorporates multi-scale models of individual cell behaviours, including subcellular networks, cell-cycle progression and cell-cell mechanics/interactions. The subcellular network allows us to incorporate our GRN, which in turn affects the cell-cycle model, eventually leading to cell-division which affects our mechanistic system forming new cells with their own subcellular networks.

We describe below, in more detail, how each of these modules is implemented in the agent-based framework for mESCs. It has been shown that cells in a NH state remain pluripotent, able to differentiate into any other type of cell, and so proliferate more quickly than those in a NL state. This suggests that cells that are found to be in a NH state have a shorter cell cycle than those in a lower state, which informs our model. We, therefore, linked the cell cycle for each of our cell agents to the state of the GRN that we have applied to our system. Two possible models

for the cell-cycle behaviour of mESCs have been suggested [83]:

a. An inherited cell cycle,

b. A dynamic cell cycle.

Considering (a) first, we assume that the length of the cell cycle for the mESCs is dependent on the level of Nanog of the mother cell at the point of mitosis (cell division). If the mother cell exhibits a high level of Nanog at mitosis, then both the daughter cells will proliferate more rapidly than if the mother cell exhibits a low level of Nanog. This initial model is based on the assumption that the fate of the cell is fixed at birth and does not subsequently (dynamically) update, although the level of Nanog within the cell is able to fluctuate and switch from a high to low state at any point within its lifetime. Figure 4.6 shows the cell-cycle length distributions for NL and NH cells; these are the distributions from which are drawn the fixed cell-cycle lengths for the cells at the point of cell division.

In contrast, the dynamic cell-cycle model (b) assumes that the fate of the cell is dynamic and updates with the current state of the cell at any point within its lifetime. Unlike (a), where each cell is allocated a fixed length cell cycle at birth depending on the state of the mother cell, in (b) each cell has a probability of division which is lower in cells which exhibit a low level of Nanog, and vice versa. Figure 4.9 shows the probability of cell division for each of the two Nanog states, with a temporally dependent logistic function governing it [83]. By considering the possibility of a dynamically updating cell fate, we can model the temporal fluctuations of Nanog and their possible effects on the cell cycle.

Considering the mechanical model of the population, we used a simple model for individual cells [120]. Assuming that each cell can be modelled as a point-centre, linked to neighbouring cells by a spring, we captured the physical adhesion of cells in a simple manner as described in Chapter 2. We do this as the mechanical considerations of the cultured populations is not likely to affect the dynamics of the system in our case, where there are optimal culture conditions without any contact inhibition or apoptosis. To take into account cell division, when a cell agent reaches the end of its cell cycle it divides; in the agent-based model, this means a new daughter agent is introduced into the simulation in a random position half a cell radius from the centre of its mother cell. This causes the rearrangement of cells surrounding the two new cells and thus affects the mechanics model. Whilst the mechanistic considerations in our agent-based model are minimal, it still forms a tool which can be extended in any future research, to include culturing beyond 4 days at which point the mechanistic set-up would become more necessary, as at this stage there is increased cell apoptosis and contact inhibition as cells form a confluent layer [68].

The agent-based models were run for 96 simulated hours, to be comparable with the experimental data (see Figure 4.12). Each of the simulations was initialised with approximately 200 cells, by the end of the simulation there were approximately 12,000 cells. By starting with this initial cell count we avoided ending the simulation with an unwieldy cell count (starting the

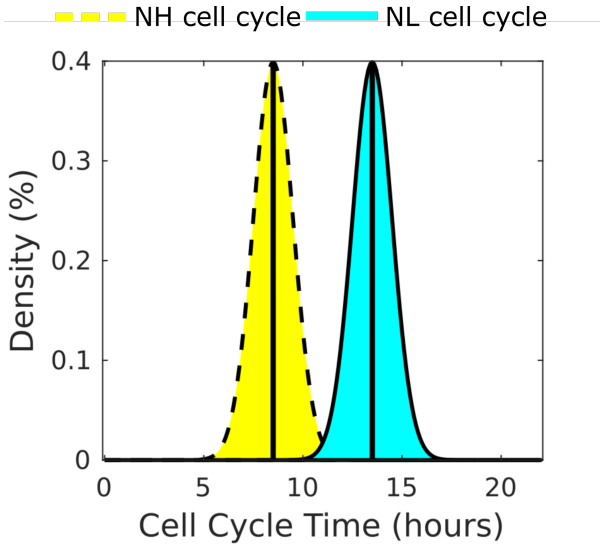


FIGURE 4.6. **Cell-cycle length distribution for both NH (yellow region) and NL (blue region) cells in agent-based simulations.** The subcellular GRN in Figure 4.3 effectively recapitulates the NH/NL distribution that is seen experimentally (an approximately 90% NH population with an approximately 10% NL subpopulation). Cells in a NL state proliferate at a slower rate than those in a NH state.

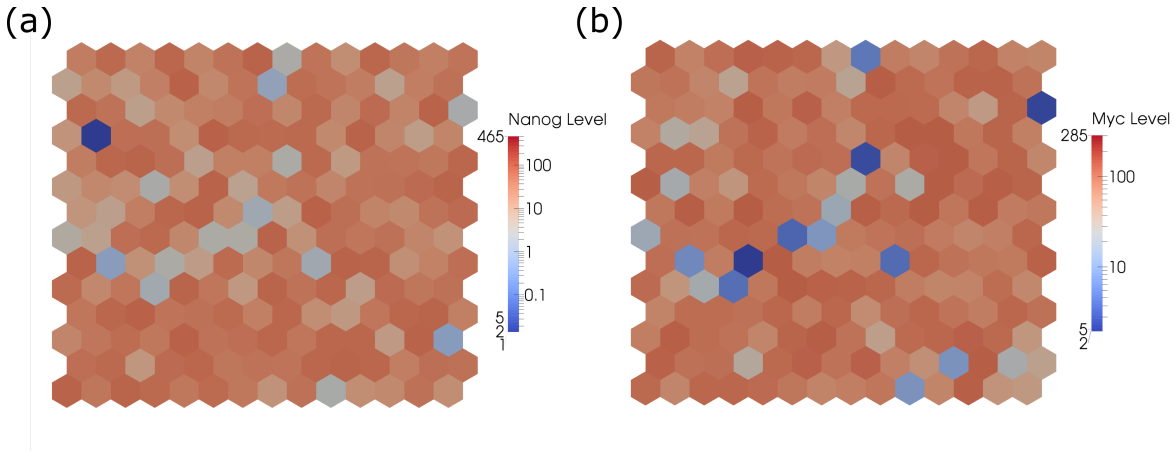


FIGURE 4.7. **Initial state of a simulated agent-based population.** Example simulation set-up of a population of mESCs in a NH starting state. Cells are considered point centres, linked to neighbouring cells by spring constraints. The cell morphology results from tessellating the population of point centres, as shown by the corresponding honeycomb placement of the cells. The colours represent the concentration of Nanog (left) and MycN (right) within the cells. The NH state shown in this example manifests as a predominately red population.

simulation with 1000 cells resulted in approximately 64,000 cells at the end). The time-step for the simulation was 1 minute (1/60h). At each time-step, the GRN was integrated using an implementation of the Euler-Maruyama method [85]. This is a simple generalization of the Euler method, used for solving ordinary differential equations (ODEs), for solving stochastic differential equations. It effectively extends the original Euler method for ODEs to include an additional random variable term, which applies an additional Wiener process (the same as standard Brownian motion process), where the random variables are independent and identically distributed normal random variables, each with a mean of zero and a variance that we prescribe (Table 4.1).

An example of the initial conditions of the agent-based simulations using cell-cycle model (a) is shown in Figure 4.7, which reproduces the initial population in a NH state (with the corresponding initial state of MycN, another gene in the extended network which, as shown in Figure 4.3, is inhibited by Nanog). Figure 4.7 (a) shows the cells in the culture in a mostly NH state (corresponding to a culture of 75% NH). It also shows on the right the corresponding levels of MycN in the cells. These are simply samples of the initial and final states of the *in silico* cell culture, as the initial conditions are changed for each simulation with individual cells initialised with cell cycles drawn from the normal distributions shown in Figure 4.6.

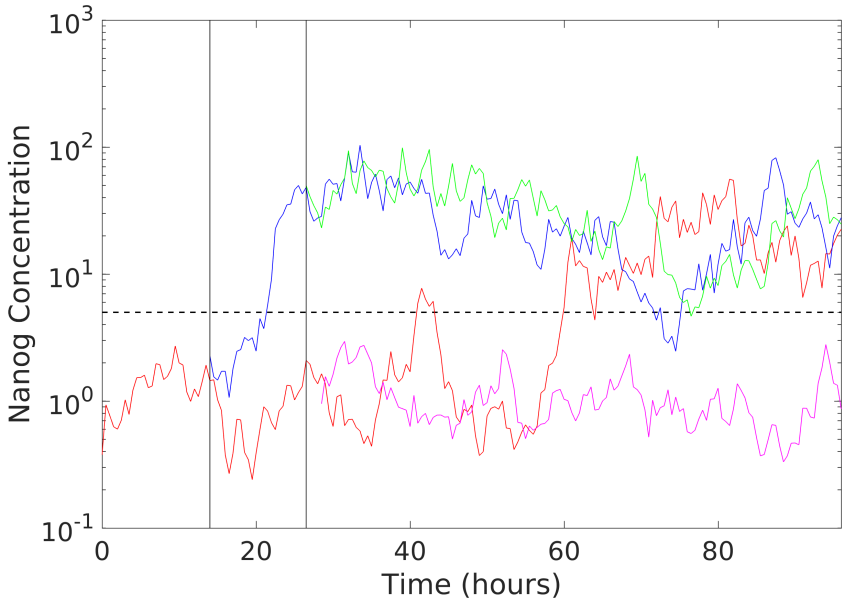


FIGURE 4.8. **Example trace of the concentration of Nanog within a single cell within the agent-based model of mESCs.** Red line corresponds to the concentration of Nanog within the initial 'grandmother' cell, with the blue line representing the first cell-division and introduction of a daughter cell, and the subsequent green and purple lines representing the second rounds of cell division for the cells. This trace provides an example of how the concentration of Nanog displays bistable dynamics, with the threshold between a NH and NL state represented by the dashed black line.

Each cell contains its GRN, shown in Figure 4.3, and corresponding SDEs in Equations 4.1 – 4.9, which describe how the concentration of each of the genes evolves in time. At each time-step, Equations 4.1–4.9 (that describe the dynamics of these gene concentrations) were solved, with cells either classified as exhibiting NH or NL states.

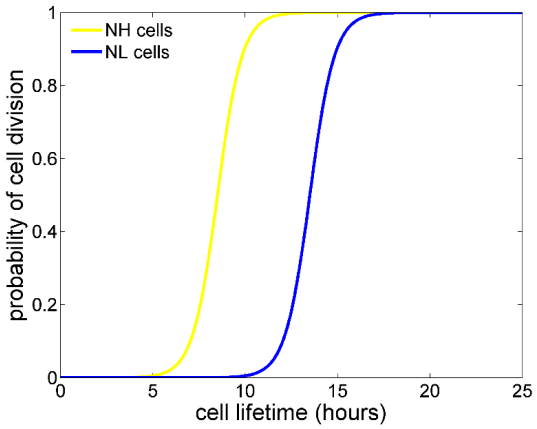


FIGURE 4.9. **Dynamically updating cell-cycle probabilities for NH/NL cells.** Cell division probabilities for both NH and NL cell states. As a cell ages, it becomes more likely to divide, with a greater probability when it is in a NH state. This dynamic function allows the fate of a cell to dynamically change during the course of a cell cycle.

Figure 4.8 shows an example of the 'inherited' model simulation, tracing Nanog levels in an individual cell. It shows that the Nanog state (whether high or low) can fluctuate over the course of the cell lifespan.

In the next section, we will analyse the results of our agent-based model. We investigated two cell-cycle models, hypothesising it is linked to the levels of Nanog (as previously suggested), or MycN; the latter gene has been shown to play an important role in the proliferative and differentiative dynamics of mESCs [150].

4.4 Coupling Nanog and MycN Levels to the Cell Cycle in Agent-Based Simulations

4.4.1 Nanog-Dependent Model

Our aim is to understand how different agent-based modelling assumptions affect the dynamics of the mESC population in different culture media. We considered four different model cases, in order to combine both versions of the cell-cycle model (inherited and dynamic) for both the serum/LIF and the 2i/LIF culture media (Table 4.2).

Set-up	Culture medium	Cell-cycle model	Dependent genes
(S1)	Serum/LIF	Inherited	Nanog
(S2)	2i/LIF	Inherited	Nanog
(S3)	Serum/LIF	Dynamic	Nanog
(S4)	2i/LIF	Dynamic	Nanog

TABLE 4.2. **Different model set-ups for the agent-based mESC system.** These different models combine the dual versions of cell-cycle model (inherited/dynamic) with the two culture media (serum/LIF and 2i/LIF).

Set-up	NH Lower (h)	NH Upper (h)	NL Lower (h)	NL Upper (h)	Interval (h)
(S1)	8	10	13	15	0.2
(S2)	12	14	20	22	0.2
(S3)	8	10	13	15	0.2
(S4)	12	14	20	22	0.2

TABLE 4.3. **Table listing the ranges and interval set-ups for the parameter sweep fitting of the cell-cycle models described in Table 4.2.** Each simulation set-up was repeated 10 times.

Table 4.2 presents the different possible set-ups for our agent-based model, with the two cell-cycle models in the different culture media. The cell-cycle parameters were fitted by running parameter sweeps across a range of biologically reasonable values. We simulated mESC culture with cell-cycle lengths according to the parameter sweep conditions in Table 4.3. We ran simulations for each cell-cycle length within the ranges in Table 4.3 10 times in an initial parameter sweep; this allowed us to narrow down search range for the optimal cell-cycle lengths. Following the initial parameter sweep, we reduced the search range from 2 hours to 1 hour and our sweep time-step to 0.1h.

Each set-up of the agent-based model was simulated 50 times, starting from different initial conditions for the genes in question drawn from normal distributions shown in Figure 4.4. Figures 4.10 and 4.11 show the resultant distributions of Nanog using the different model set-ups (Table 4.2). Table 4.4 shows the fitted cell-cycle parameters for the Nanog dependent cell-cycle model in both Serum/LIF and 2i/LIF. The cell-cycle times for the 'inherited' model (S1 and S2) correspond to the mean of a normal distribution ($N(NH/NL, 1)$, with mean NH/NL and variance 1), as shown in Figure 4.6. For the 'dynamic' cell-cycle model, the times refer to the mid-point of a logistic probability function which determines the probability of a cell dividing at a given age, as shown in Figure 4.9.

The initial results from set-ups S1 to S4 suggested that the cell-cycle length is extended when culturing cells in 2i/LIF compared to serum/LIF, with the mean length of the NH cell-cycle is 55.3% longer in 2i/LIF. The mean length of the NL cell-cycle is 54.8% longer in 2i/LIF compared to

Set-up	Nanog noise (σN)	NH cycle length (h)	NL cycle length
(S1)	0.44	8.50	13.5
(S2)	0.35	13.2	20.9
(S3)	0.44	8.20	13.2
(S4)	0.35	12.7	20.5

TABLE 4.4. **Table listing the parameters fitted for set-ups 1-4 of the agent-based simulations.** The parameters include the stochastic differential equations (SDE) noise term, taken from a normal distribution $N(0,\sigma^2)$, and the cell-cycle lengths for the corresponding cell-cycles. The cell-cycle models are dependent on the state of Nanog within the cell. For the inherited cell-cycle, the cell-cycle lengths are allocated from a normal distribution ($N(NH/NL,1)$ with mean NH/NL and variance 1). For the dynamic cell cycle, the cell-cycle length determines the midpoint of a logistic probability function, as shown in Figure 4.9.

serum/LIF. In order to analyse how the system has changed between these different model cases, we plotted the distribution of Nanog within the population and compared how they changed over time. In the case of the populations cultured in serum/LIF (Figure 4.10), all of the populations remain close to a 50% NH ratio, corroborated by the distributions shown. The 2i/LIF culture simulations shown in Figure 4.11 similarly show that the populations tend towards the desired ratio of NH to NL (90%NH/10%NL).

Whilst these models suggest that we are able to match to the required steady-state distributions of Nanog within the population of mESCs with all 'inherited' cell-cycle set-ups in both serum/LIF and 2i/LIF, the results suggest that there is not a significant difference between the 'inherited' and 'dynamic' cell-cycle models in terms of matching the distributions, however, the dynamic model of the cell-cycle is computationally more intensive due to random number generation required for each cell in the simulation at each time-step. As both the 'inherited' and 'dynamic' cell-cycle models capitulated the required experimental data, in what follows we simply considered the inherited model.

We then collected time-course experimental data, measuring mESC proliferation in the two different culture media; in Figure 4.12, the solid black line depicts the fold change for serum/LIF culture and the dashed black line the fold change for the 2i/LIF culture experiments, with the standard deviation for the experiments (across three replicates) shown as the shaded grey regions. The cell counts were measured at 24-hour intervals, over a time-course of 72 hours.

Figure 4.12(a) shows the resultant fold change captured by our agent-based model (S1/S2) when we fit the model to the serum/LIF data (blue line) with the corresponding standard deviation shown as the shaded blue region. Using the cell-cycle parameters fitted (shown in Table 4.4 S1 & S2) to the serum/LIF experimental data we can then simulate the effects of switching our model to a 2i/LIF culture medium (i.e. modifying the PD and CH values in the SDE single cell models). In 2i/LIF medium we measured a reduction in the cell count fold change, but in our agent-based

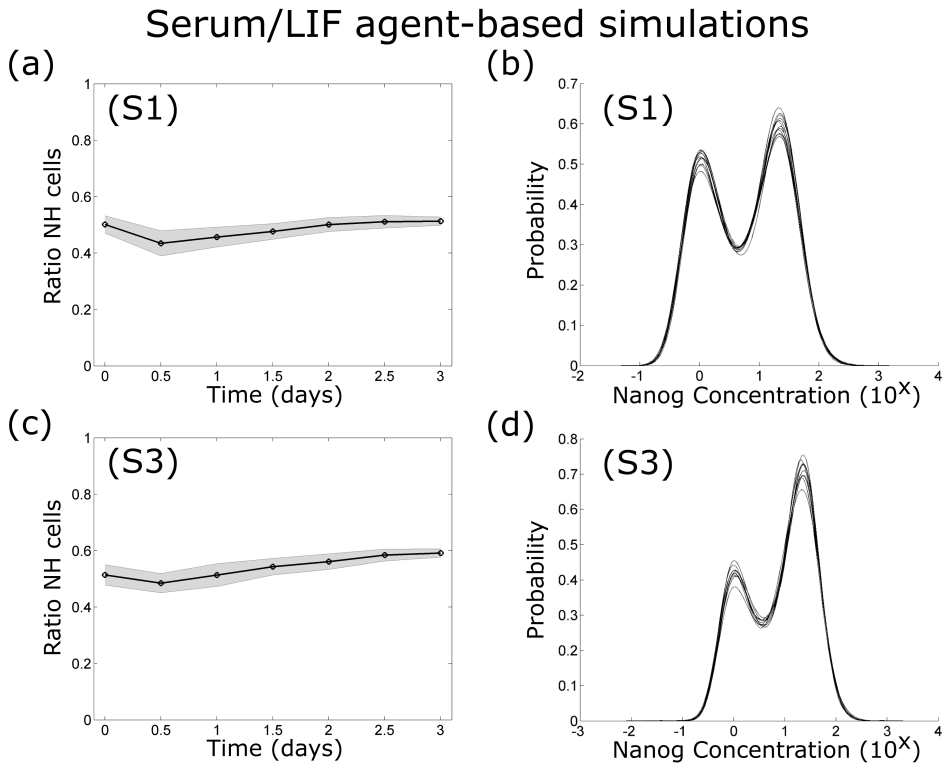


FIGURE 4.10. **Distributions of Nanog within the population of mESCs cultured in serum/LIF from agent-based models as in Table 4.2.** Ratio of NH to NL cells (a,c) for the different cases, and the Nanog concentration distributions (b,d).

model *in-silico* experiments we predict an increase in the fold change compared to the serum/LIF medium and significantly greater than the experimentally expected fold change (more than 4x the fold change as shown in Figure 4.12 comparing the red line (simulation) with the solid black line (experimental)), suggesting that our simulated cell population is proliferating too rapidly. Alternatively, we fitted the agent-based model to the experimental data for 2i/LIF, shown in Figure 4.12(b). The solid red line shows the fitted 2i/LIF model predictions, with the blue line showing the predicted fold change for our simulated serum/LIF culture; again, we observe a significant overshoot in the fitted serum/LIF, and a significant undershoot in the fitted 2i/LIF. This suggests that coupling the cell-cycle length to concentrations of Nanog might not be effective in replicating the experimentally captured proliferation dynamics.

4.4.2 MycN-Dependent Model

Whilst we were able to capture some of the required population dynamics with the Nanog dependent model, they were unable to capture the population count changes seen when changing from serum/LIF to 2i/LIF. This caused a large overshoot in cell-count within the computational model population. One potential fix for this issue is to include a different gene dependence in

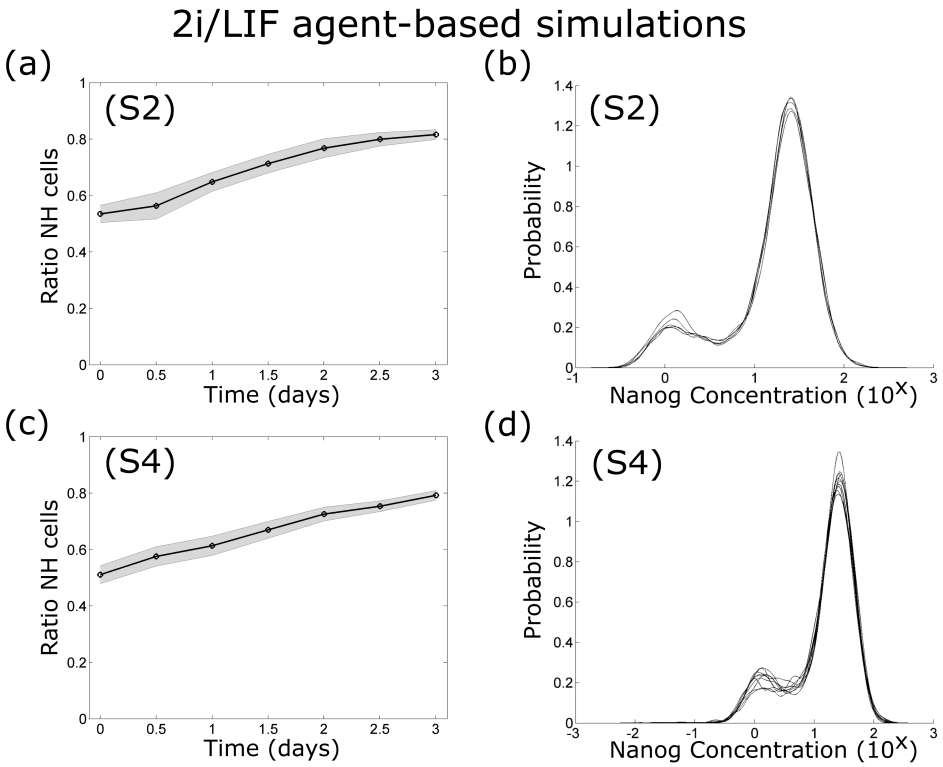


FIGURE 4.11. **Distributions of Nanog within the population of mESCs cultured in 2i/LIF from agent-based models as in Table 4.2.** Ratio of NH to NL cells (a,c) for the different cases, and the Nanog concentration distributions in the population (b,d).

our cell-cycle model. Figure 4.13 shows the distribution of MycN within the mESC population, described by the extended network for the two different culture media from the same GRN network solutions shown in Figure 4.4. The blue and orange lines show the expected MycN distribution in 2i/LIF and in serum/LIF, respectively. As with Nanog in the extended network, switching between the two culture media results in a shift in the distribution of MycN high (MH) and MycN low (ML) cells [114], with two peaks shown, but slightly less pronounced than those in the Nanog distribution in Figure 4.4.

In order to use MycN as the basis for alternating the cell-cycle model, we considered its effect on the cell cycle in the same way as Nanog. As the differences between the 'inherited' and 'dynamic' cell cycles were small, we solely considered the inherited cell-cycle model. We allocated normally distributed cell-cycle lengths, dependent on whether cells are in a MH or ML state at the point where a mother cell divides. This allocation to a MH or ML state is accomplished by creating a threshold at the midpoint of the peaks in the MycN distributions shown in Figure 4.13. To fit the MycN cell-cycle lengths, we applied the same methodology that we used in Section 4.4.1. We iterated over a range of cell-cycle lengths to best fit 4.4, the resultant cell-cycle lengths for

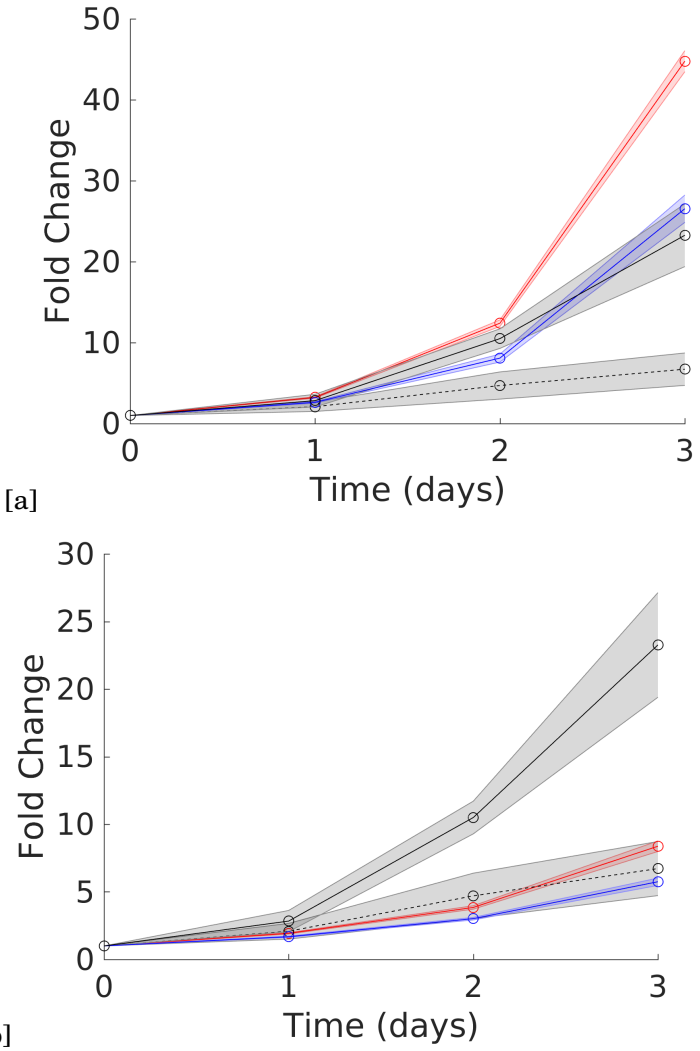


FIGURE 4.12. **Comparison between experimental data and simulated cell population with Nanog dependent cell cycle.** (a) Cell-cycle model fitted to serum/LIF data (blue line represents simulated cell count fold-change, solid black line represents serum/LIF experimental data) and then used to predict the fold change in a 2i/LIF culture (red line represents the predicted 2i/LIF fold-change, black dashed line represents 2i/LIF experimental data). (b) Cell-cycle model fitted to 2i/LIF data (red line represents simulated cell count fold-change, solid black line represents serum/LIF experimental data) and then used to predict the fold change in a 2i/LIF culture (blue line represents the predicted 2i/LIF fold-change, black dashed line represents 2i/LIF experimental data).

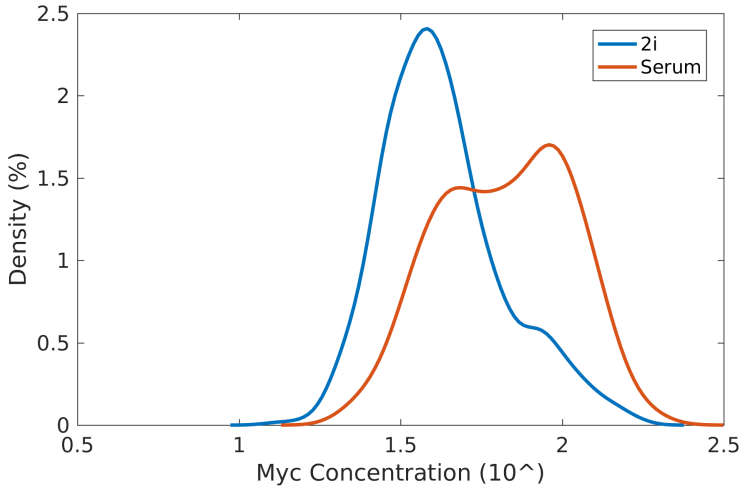


FIGURE 4.13. **Distribution of MycN concentration.** Simulated distributions of MycN concentration within a mESC population when cultured in serum/LIF or 2i/LIF, showing the switch from a majority MycN high distribution in serum to a majority MycN low distribution in 2i/LIF. This difference in expected population distribution between the two media will affect the cell-cycle when linked to the distributions in Figure 4.6. We can use the differential distributions in MycN concentration as the link between the cell cycle and the GRN to better fit the experimental data.

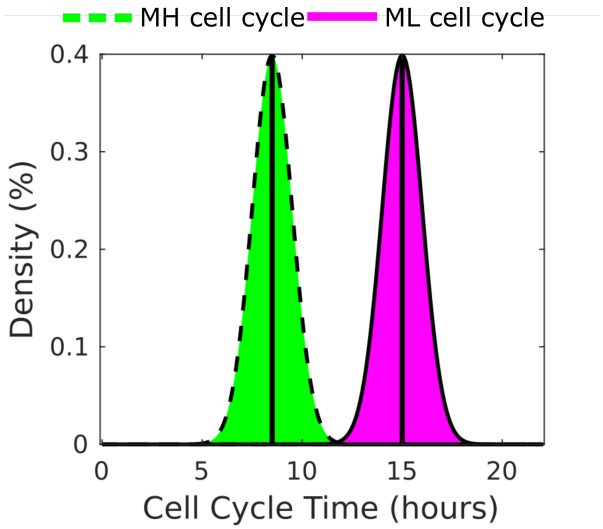


FIGURE 4.14. **Cell-cycle length distributions for the MycN dependent cell cycle.** The cell-cycle length is sampled from a normal distribution, which is dependent on the MycN concentration within the cell. MycN low cells are allocated a cell-cycle from the green distribution (mean: 8.5h, std: 1h), and MycN high cells are allocated from the purple distribution (mean: 15.0h, std: 1h)

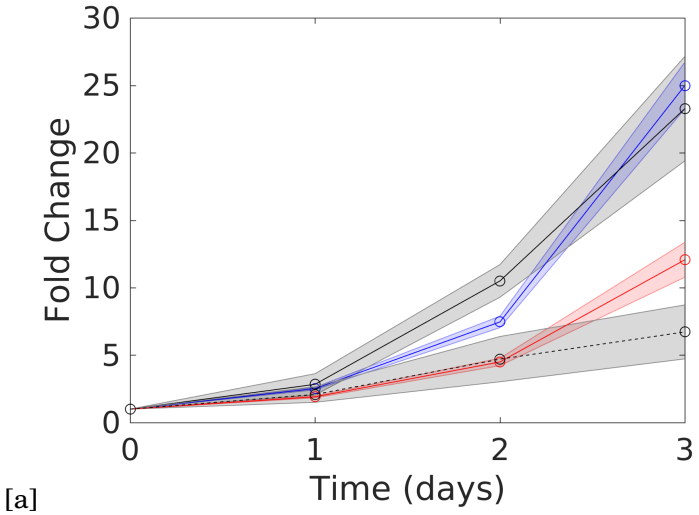
Set-up	Nanog noise (σN)	MH cycle length (h)	ML cycle length
(S5)	0.44	8.5	15.0
(S6)	0.35	9.5	16.0

TABLE 4.5. **Table listing the parameters fitted for set-ups 5-6 of the agent-based simulations.** The parameters include the stochastic differential equations (SDE) noise term, taken from a normal distribution $N(0,\sigma^2)$, and the cell-cycle lengths for the corresponding cell-cycles. The cell-cycle models are dependent on the state of MycN within the cell. The cell-cycle lengths are allocated from a normal distribution ($N(MH/ML,1)$ with mean MH/ML and variance 1).

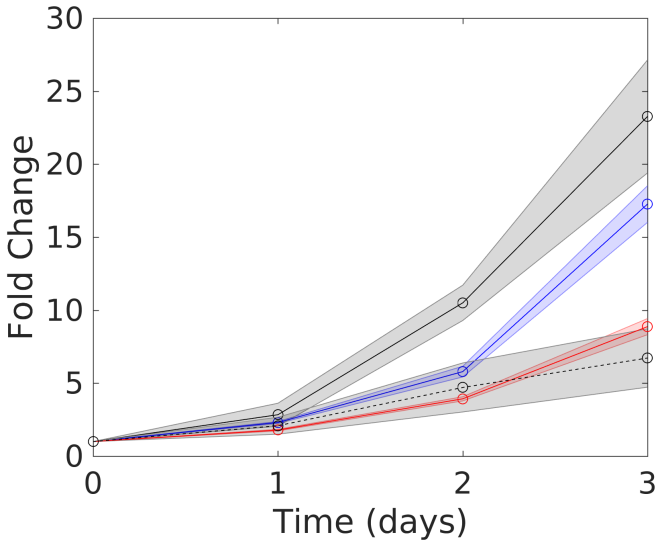
S5 (MycN dependent cell-cycle model in Serum/LIF) and S6 (MycN dependent cell-cycle model in 2i/LIF). Similar to our Nanog cell-cycle dependence [114], Figure 4.14 shows the resultant cell-cycle length distributions for the two MycN states, MH (purple distribution) and ML (green distribution) with means of 15.0h and 8.5h respectively and both with standard deviations of 1h. These distributions are fitted such that cells that are in a MH state proliferate more rapidly than those in a ML state, following a similar parameter sweep scheme as that shown in Tables 4.3. This modelling assumption is motivated by recent experimental data [150].

Using MycN as the sole cell-cycle differentiator for our mESCs, we fitted our model to the experimental data for serum and predicting the fold change in 2i/LIF, and then fitting the experimental data for 2i/LIF and attempting to predict the fold change for serum/LIF. Figure 4.15(a) shows the results from our agent-based model (blue line) for fitting to the serum/LIF experiments (solid black line). We were able to qualitatively capture the fold change seen in the culture of mESCs in serum/LIF by using the MycN dependent cell-cycle. Switching our agent-based model to 2i/LIF medium by modifying PD and CH in the GRN we predicted the fold change seen in 2i/LIF medium. The experimental data for the 2i/LIF culture is shown as the dashed black line, the predicted fold change seen in the agent-based model (red line) better captures the fold change, but most importantly it shows that switching to this medium reduces the fold change in the cell counts whereas using Nanog we get a significant increase in the fold change. The results of attempting to do the opposite, i.e. fitting the agent-based model to the 2i/LIF experimental data and predicting the serum/LIF results, are shown in Figure 4.15(b). We were again able to qualitatively replicate the fold change seen in the experimental cell culture, shown as the red line. Switching back to serum/LIF with the 2i/LIF fitted parameters resulted in an increase in the fold change seen, as expected.

Finally, to test whether our agent-based model accurately captures the dynamics we expect from the single cell extended network experiments, we tested how the shift in NH proportion changes over time. Upon sorting, in serum/LIF culture it is expected to see an approximately 50%/50% split in NH and NL cells, while in 2i/LIF it is shifted towards 75%/25% NH/NL [83]. In Figure 4.16(a) we show how the distribution of NH/NL cells changes within the culture: starting from a fully NH population (red line) we see a shift towards the 50%/50% distribution expected,



[a]



[b]

FIGURE 4.15. **Comparison between experimental data and simulated cell population with MycN dependent cell cycle.** (a) Cell-cycle model fitted to serum/LIF data (blue line represents simulated cell count fold-change, solid black line represents serum/LIF experimental data) and then used to predict the fold change in a 2i/LIF culture (red line represents the predicted 2i/LIF fold-change, black dashed line represents 2i/LIF experimental data). (b) Cell-cycle model fitted to 2i/LIF data (red line represents simulated cell count fold-change, solid black line represents serum/LIF experimental data) and then used to predict the fold change in a 2i/LIF culture (blue line represents the predicted 2i/LIF fold-change, black dashed line represents 2i/LIF experimental data).

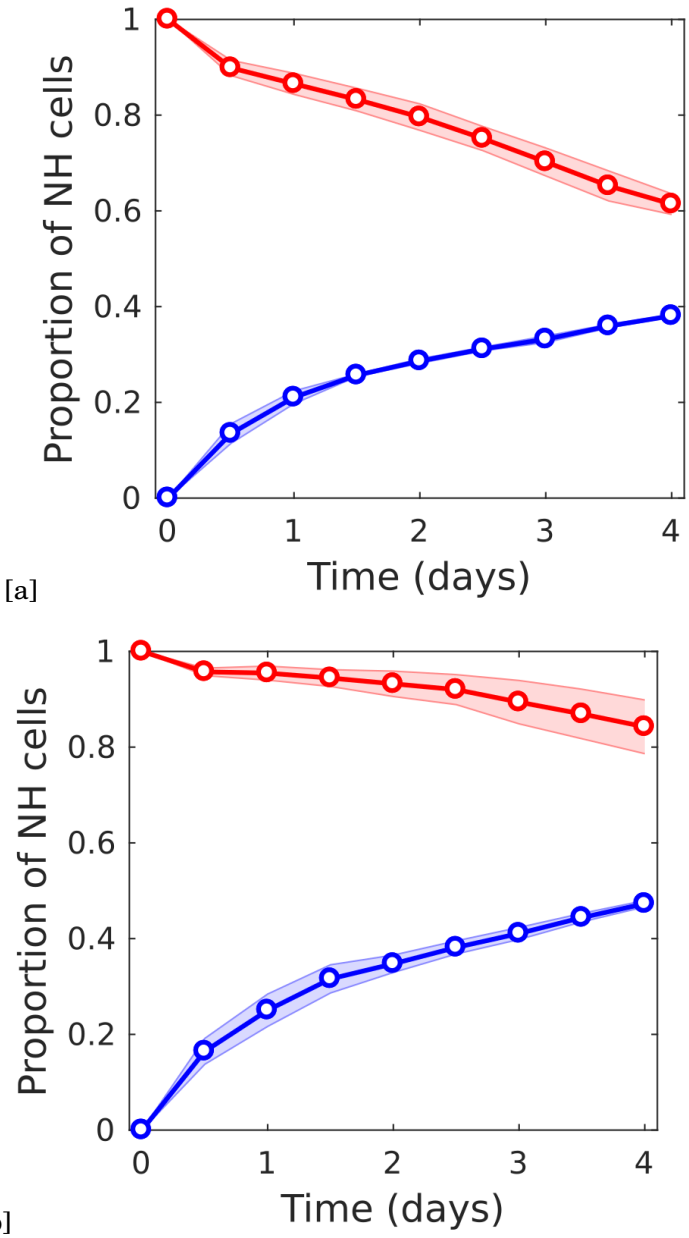


FIGURE 4.16. **Results from *in silico* cell sorting experiments in serum/LIF (a) and 2i/LIF (b).** Starting from purely NH/NL states, we simulate the system for 96h and track the ratio of cells in a NH state over time. The red line depicts the changing ratio of NH cells in a population starting from a NH state and the blue line depicts the changing ratio of NH cells from an initial NL state, for simulated serum/LIF (a) or 2i/LIF (b) cultures. The resultant ratios are averages across 15 independent simulations, with the shaded region showing the standard error.

as well as when starting from a fully NL population. This suggests that our agent-based model, with gene-differentiated cell-cycle lengths (and thus gene-dependent proliferative capacity), is able to quantitatively capture the dynamics of the sorting experiments performed in Herberg *et al.* [83]. Figure 4.16(b) shows the results of modelling the sorting experiments in a 2i/LIF culture. Again, starting from a NH population we predict a shift towards the expected 75%/25% distribution (red line), with a more rapid shift when starting from a NL population (blue line).

4.5 Discussion

In this chapter, we have described a novel extended gene-regulatory network (GRN) for mouse embryonic stem cell (mESC) dynamics [68], designed such that it can replicate the Nanog dynamics both in terms of their bistability [2, 3] and population distribution. A SDE system models this novel extended GRN and captures the experimental distributions of Nanog measured in both the serum/LIF and 2i/LIF media. To determine whether the results of the single-cell analysis of this extended network applied at a population level, we developed an agent-based model of mESC culture, incorporating the GRN to capture internal dynamics, and linked it to additional biological processes. We linked the state of the underlying GRN to the cell cycle, whether a cell exhibits high or low levels of key genes (Nanog/MycN), to investigate any emergent population dynamics from differential cell-cycle lengths affecting population growth. We investigated the role of Nanog in regulating cell-cycle behaviours, as suggested by previous studies [83], but additionally, we extended this investigation to the role of MycN, an alternative gene whose role in regulating cell proliferation and differentiation is less thoroughly studied but potentially significant [150].

By using Nanog as the cell-cycle controller in our agent-based model, as suggested in previous work [83], we were able to quantitatively capture the distribution of Nanog seen experimentally, comparing Figure 4.1 and Figure 4.10(b). However, when we considered the population growth dynamics, something that we can only do using our agent-based model, there results a significant (80% greater than expected) overestimation of the population count of the cells when we switched from the serum/LIF medium to 2i/LIF, as shown by Figure 4.12(a). This overestimation when switching culture medium is not something seen in the previous studies, which focussed solely on the serum/LIF culture condition. In order to capture both the gene concentration distributions as well as the growth dynamics, we considered the alternative gene, MycN, as the regulator of cell-cycle behaviour. By linking the concentration of MycN present within mESCs to their proliferative capacity we were able to capture the fold change characteristics of the cell culture in both media, as well as capturing the distribution of Nanog within the population. Using MycN as the differentiator or cell-cycle behaviour has not, as far as we are aware, been applied before, with Nanog seen as the controller of proliferation in these pluripotent cells. Our fitted agent-based model was able to produce simulated dynamics that matched qualitatively (by capturing the expected reduced proliferation in the switch from serum/LIF to 2i/LIF) experimental data, as shown

in Figure 4.15, and quantitatively (capturing the distributions expected from the experiment shown in Figure 4.16). Additionally, we were able to reproduce the sorting experiment results simulated using the single cell SDE model, with the *in-silico* sorting experiment results in Figure 4.15(a) tending towards the expected 50%/50% distribution, and Figure 4.15(b) tending towards the expected 90%/10% distribution [2, 3]. Considering the effects of MycN on the proliferative behaviour of the cells has allowed us to shine some light on the chemical dependencies of these highly important cells, extending our knowledge on control and maintenance of mESCs *in vitro*.

Applying agent-based models to these populations allowed us to analyse the effects of different scales within the population, however, we have had to make simplification and assumptions to our model. This work could be extended to consider the implications of cell-cell interactions within the population, in order to consider effects such as contact inhibition and cell apoptosis. Additionally, by extending this model to consider non-confluent populations of cells or cells able to extend within the domain, you can consider spatial rippling and patterning and the effects this may have on the propagation of cells carrying high or low levels of these factors within the domain. Although the exclusion of physical factors did not detract from the modelling, and the applicability to the experimental data, the consideration of additional features could yield additional interesting dynamics.

Whilst we have shown in this chapter that the use of agent-based models allows us to incorporate single-cell systems, namely an underlying gene-regulatory network, within a population level framework in order to validate the single-cell analysis whilst simultaneously investigating possible emergent population level dynamics, it is possible to increase the complexity of these models for different cell populations. In the following chapter, we will consider agent-based modelling of increased complexity, incorporating a more extensive subcellular network, linked to a bistable ODE representation of cell-cycle progression. This extends on the prescribed cell-cycle model considered in this chapter, providing an opportunity to analyse whether the increased complexity leads to an increased biological relevance. In addition, we consider cell-level physical interactions that play a key role in the dynamics of the population by developing a model for contact-inhibition.

Chapter 5

Multiscale Modelling of Intestinal Crypt Dynamics

Intestinal crypts are testubular indentations that line the walls of both the small and large intestines, containing the stem cells which are key to the renewal of the intestinal lining every 4-7 days [27]. Although the healthy maintenance of the intestinal wall due to the renewal of this lining typically results in the expulsion of possible mutant cells, there still exists the risk of dysplasia to the healthy formation of the crypt, and subsequent spreading of mutant cells to surrounding crypts along the wall of the intestines [10].

This chapter extends on the modelling techniques in Chapter 4, which used agent-based modelling methods to replicate the growth dynamics in different culture conditions of mouse embryonic stem cells (mESCs), by adding extended spatial model complexity. We both coupled cell-cycle dynamics to internal levels of specific genes and signalling pathways, and considered the physical interaction of cells within crypts, by implementing the biological process of contact inhibition for intestinal stem cells (ISCs). Thus, unlike the agent-based modelling in Chapter 4, ISCs cell cycles are dependent on both the internal state of their subcellular network and their physical interactions within the cell population. We investigated the role of the ISCs residing at the base of the crypt in maintaining crypt homeostasis, and possible causes for dysplastic dynamics that have the capacity to transform a healthy crypt into an unhealthy mutant crypt. Our analysis shows that contact inhibition can plausibly reduce the proliferative capacity of cells within intestinal crypts, as a reduction in individual cell proliferation causes a global reduction in crypt renewal; also, it is possible to link the onset of contact inhibition to signalling pathway dynamics. We extend on a previous single cell network model describing the subcellular localisation of β-catenin (the main effector of the Wnt/β-catenin signalling pathway) within

the cell and its effect on governing cell proliferation by incorporating a mechanism for Hippo signalling dependent contact inhibition. This mechanism involves linking the physical state of the cell to its level of Hippo signalling, which in turn affects the localisation of β-catenin, with increased Hippo signalling causing a reduction in both β-catenin nuclear localisation and transcription of genes promoting cell-cycle progression. We also simulated mutations in the Wnt-signalling pathway, affecting β-catenin and also Hippo signalling; this impacted on the probability of a mutant cell taking over a healthy crypt.

In Section 5.1 we provide the motivation for this work, developing tools which allow us to better understand the onset of crypt dysplasia. In Section 5.2 we provide the background biology related to the intestinal crypt system, describing the physical setup including the various cell types and external influences to the population, as well as the biological processes we have modelled. In Section 5.3 we present an initial analysis of a basic intestinal crypt model implementation of cellular contact inhibition. In Section 5.4 we model the cross-talk between Wnt and Hippo signalling, and subsequently implement this model in an agent-based framework to investigate emergent dynamics. Finally, in Section 5.5 we discuss the results from our computational investigation into ISCs and the role they play in the healthy homeostasis of the intestinal crypt, as well as the role of mutants in the onset of crypt dysplasia.

5.1 Motivation

Colorectal cancer (CRC) is the third most common cancer in the UK [185], with over 37,500 people diagnosed each year. Although the overall mortality rate for colorectal cancer has been declining by roughly 2% per year between 1997 and 2007 in the EU, from 19.7 to 17.4/100,000 men and from 12.5 to 10.5/100,000 women [17], mainly due to improved early diagnosis and/or an improved lifestyle, colorectal cancer still affects a significant number of people worldwide. Improvement of these statistics (through, for example, new diagnostics) will depend on a better understanding of the cellular mechanisms that are responsible for the disease. Although it has been suggested that there are specific cancer stem cells that develop and propagate colorectal cancer [63, 111], this is not fully understood and there are still suggestions that the propagation of cancers within the intestines begins with individual mutations to originally healthy cells residing at the base of the intestinal crypts. In either case, there exist several key biological questions that need answering in order to further understand what is happening during the initial stage of colorectal cancer. These questions can be considered across several physiological scales, from trying to understand causes on a genetic subcellular level, in terms of mutation dependent cancer propagation, to whether the geometry and cellular interactions have an effect on how cancer spreads.

Studies have investigated specific signalling pathways involved in crypt dysplasia (e.g. the Wnt pathway [91, 107, 172], the Notch pathway [15, 22, 174]), focusing on the role of cell-cell

interactions in intestinal epithelium homeostasis [7, 47] as well as investigating cell migration in healthy and dysplastic crypts [48, 120, 124]. Further to this, there has been an increase in computational works providing multiscale models of the intestines, to enable a better understanding of the interplay between signalling pathways, cell cycle dynamics and spatial interactions between cells [46, 48, 49, 105, 173].

However, although several attempts have been made to model this dynamical system, key simplifications have been made in each of the models produced to date. These simplifications include reducing the complexity of the cell cycle and its link to the underlying Wnt regulatory system, and simplifications of the geometries involved. All models have so far focused on modelling single crypts in isolation, forgoing any interaction or effects gathered by having a more realistic multi-crypt system. In addition, although the role of Wnt signalling in governing cell proliferation and differentiation within the crypt has been extensively explored, recent studies have raised questions about how the Wnt gradient is formed and affects the intestinal epithelial cells. The discovery of additional pathways, such as the Hippo signalling pathway, and their role in the intestines is more recent still and thus less thoroughly understood.

In this chapter, we study the possible roles of Wnt within the crypt, in light of recent studies of the canonical Wnt signalling pathway [56], and investigate possible roles of Hippo in mediating the underlying Wnt signalling in additional biological processes such as cellular contact inhibition.

5.2 Background Biology

This section will provide the biological context and information about relevant processes involved in CRC. We provide an overview of the intestines, the intestinal epithelium, villi and crypts, and then focus on the relevant regulatory networks.

5.2.1 The Intestines

In simplified terms, the human intestinal system comprises 3 main areas: the stomach, the small intestine and the large intestine (colon), as seen in Figure 5.1. These sections of the intestinal system can then be split up into specific regions existing to undertake key tasks, such as the duodenum, jejunum and ileum in the small intestine and the ascending, transverse and descending colons within the large intestine. The focus will initially be on both these small and large intestine systems, comparing them and deciding where to investigate further. The small intestine and the large intestine both have very specific jobs; this means that, although they have some similarities, there are also key differences.

The small intestine primarily extracts and absorbs nutrients from the food we consume [54], being the first port of call of the food after it is initially digested within the stomach. The partially digested food is transported to the duodenum where it continues to be digested and the nutrients absorbed into the intestinal wall [157]. To optimise this process, the small intestine

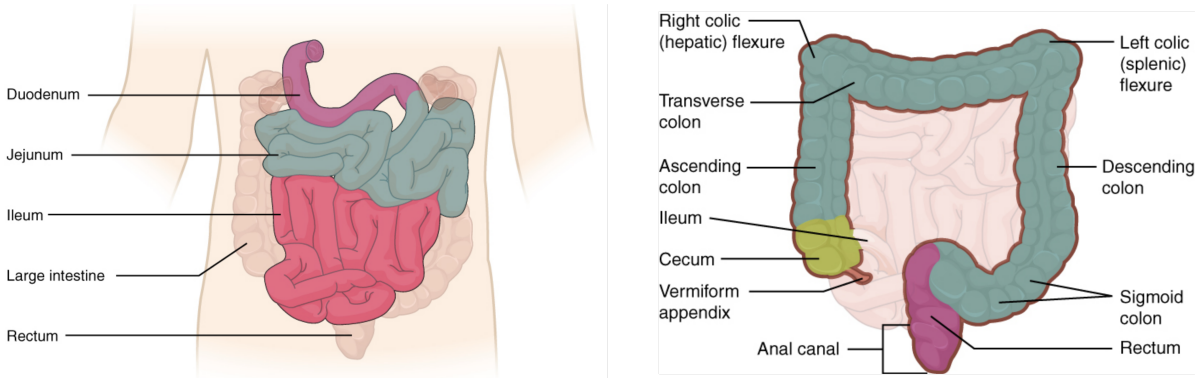


FIGURE 5.1. **Diagram showing the general layout of the human intestinal system** [13]. The key areas we are interested in are the large intestine (colon) and the small intestine

is covered in tiny finger-like protrusions called villi. These have been extensively looked at as a means by which to increase the surface area of the intestine [87], but it has also been suggested that they act to improve fluid flow within the intestinal lumen (the cavity through which food is transported), in order to optimise the number of nutrients that can make contact with the intestinal wall for absorption [42].

Unlike the small intestine, the main function of the large intestine is to absorb water from the remaining indigestible food matter. This means that, unlike the small intestine, the large intestine does not need villi lining the walls, as nutrient absorption is minimal at that stage in the digestive system. The absorptive characteristics of the large intestine, with a relatively flat epithelial wall devoid of villi, is capable of absorbing any remaining water [35].

Although the jobs of the small and large intestines within the digestive system are distinct and different, they share a key system: both contain intestinal crypts, sunken regions present along the intestinal wall which contain the proliferative system that governs total intestinal cell renewal. These crypts contain the intestines' proliferative cells, which act to replace any damaged cells, and are capable of reproducing so rapidly that the epithelial wall of the intestines is replaced every 4-7 days [27]. Although both intestinal regions contain these crypts, there are some differences in the cell types present.

5.2.2 Intestinal Crypts

The intestinal crypts are the factories of the intestines; as the intestinal epithelium is the most rapidly self-regenerating tissue of the body [27], they form a system through which all cells within the intestines are replaced and renewed. Those found in the small and large intestines do contain cells common to both, including proliferative (progenitor) cells such as stems cells and

differentiated (transit amplifying) cells. It is possible to distinguish four types of cells in the small intestine: stem cells (SCs), semi-differentiated transit amplifying cells (villus columnar cells), differentiated cells that undergo vertical crypt transport (e.g mucous and enteroendocrine cells) and finally differentiated cells that move towards the crypt base (Paneth cells [10]), as shown in Figure 5.2.

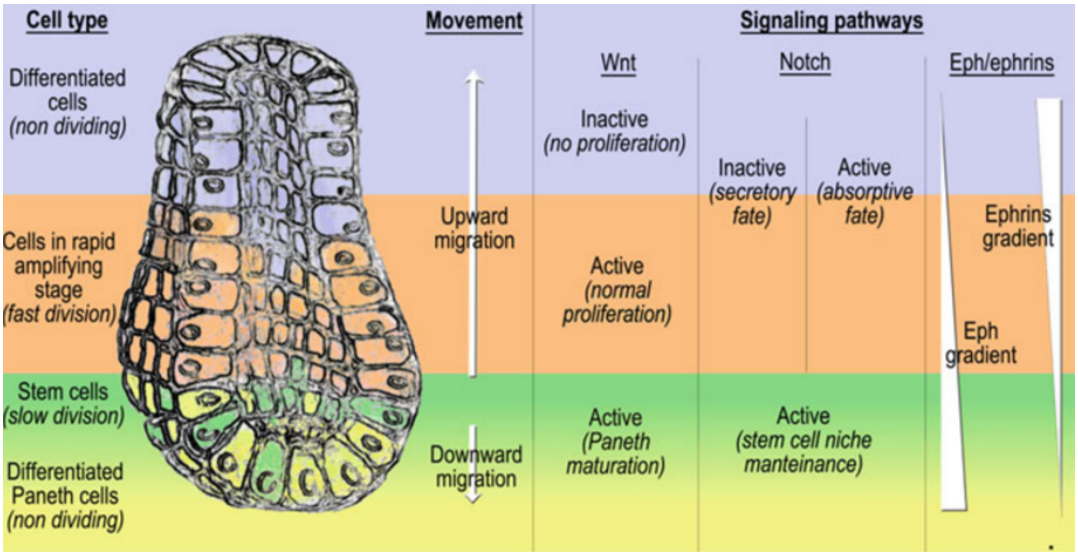


FIGURE 5.2. **Diagram showing the spatial distribution of cell types within a small intestinal crypt.** There are four main types of cell found within intestinal crypts crypts, differentiated cells (found towards the top of the crypt), transit amplifying cells (found in the middle of the crypt), stem cells (located towards the lower half of the crypt) and differentiated Paneth cells (located at the base of the crypt). The small intestinal crypt differs from the colonic crypt due to the presence of Paneth cells at the base. Key signalling pathways (Wnt and Notch) within the crypt and their relative spatial activation are also shown. Figure reproduced from Barker *et al* [10]

Figure 5.2 provides a general layout of how various different cell types are distributed in the small intestinal crypt under normal conditions. In a healthy crypt, cells towards the base of the crypt (the stem cells) reside in a favourable Wnt region proliferate and subsequently form a sudo cellular conveyor belt system which transports semi-differentiated cells towards the top of the crypt [10]. At this point, they have terminally differentiated and are then either shed into the intestinal lumen (more prevalent in the large intestine) or contribute to the formation of the intestinal villi (in the small intestine) [76]. Although the general layout of the crypts within the small and large intestine are similar, due to the differing roles each play in the digestive system there are also key differences. The most notable difference is the presence of a differentiated cell type (Paneth cells) at the base of small intestinal crypts [174]. One of the roles of Paneth cells is to synthesize and secrete substantial quantities of antimicrobial peptides and proteins

[170]. This provides protection to the epithelial stem cells from any bacteria and other damaging substances in the small intestine. Additionally, Paneth cells have been shown to be a source of extracellular Wnt within the crypt [57], as shown in Figure 5.2. Paneth cells are also the only differentiated cells within the intestinal crypt that undergo downward transit. Whilst they are present within the small intestinal crypt, they do not exist in the large intestine. This is partly because, at that stage of the digestive system, they are not required to provide additional Wnt to the crypt or to protect the epithelial stem cells, as the first line of defence in the small intestine is enough.

The signalling pathways eluded to in Figure 5.2, namely the Wnt and Notch signalling pathways, govern the proliferation and differentiation of the crypt cells. There exists a complex interplay between multiple signalling pathways [107], with the signalling of the Wnt and Notch pathways, in particular, controlling key behaviours including cell-cycle transition events and cell-lineage differentiation. In addition, there exist other signalling pathways, such as the Hippo signalling pathway, to govern cell proliferation; so these key signalling pathways will be discussed in further detail in Section 5.2.3.

5.2.3 Intracellular Signalling in the Crypt

As mentioned in the previous section, there are several types of cells present within the intestinal crypts. These cells, whether proliferating, differentiating or fully differentiated, have distinct roles in the process of epithelial renewal. These roles are controlled by underlying regulatory networks within the cells that combine cell level systems (such as cell cycle or cell adhesion) and crypt level systems (such as signalling pathways). In a healthy crypt, these systems work together to maintain stable cell counts for each of the different cell types. However, in unhealthy crypts, such as those associated to CRCs, there are changes to these underlying regulatory networks that cause the system to be disrupted, and the crypt can be taken over by unhealthy cells.

The role that the Wingless/Int (Wnt) signalling pathway plays in regulating the expression of target genes linked to cellular processes, including cell proliferation and differentiation, has been extensively studied [91, 142, 156]. Within the gastrointestinal tract, it is known that Wnt signalling governs proliferative dynamics specific to the lower regions of intestinal crypt [64]. The importance of the Wnt pathway in maintaining a healthy cell turnover is clear from studies reporting consequences of Wnt signalling dysfunction [31, 140, 156, 171], with the development of intestinal tumours and polyps linked to mutations in the signalling pathway cascade, and approximately 80% of all human colon tumours reported to show mutations inactivating the APC gene [70].

The main cellular effect of the canonical Wnt (i.e. Wnt/β-catenin) signalling pathway activation is to alter the localisation of β-catenin, the main pathway effector [31, 156]. The absence of a Wnt-signal (Wnt-Off state) leads to the rapid degradation of cytoplasmic β-catenin by the destruction complex, which is formed by adenomatous polyposis coli (APC), Axin, Ser/Thr kinases,

glycogen synthase kinase 3 (GSK-3) and casein kinase 11 (CK11) [156]. When Wnt signal is present (a Wnt-On state), extracellular Wnt-protein ligands bind to Frizzled surface receptors of the cell which in turn causes the inhibition of the destruction complex via sequestering these proteins to the cell membrane [148]. This leads to the accumulation of unphosphorylated β-catenin in the cell cytoplasm and thus increases the probability of β-catenin localising to the nucleus, where it binds to TCF/LEF transcription factors, eventually leading to the increased expression of numerous Wnt target genes [142]. Typically, malignant transformations of cells involve mutations in the APC gene [70] resulting in dysfunction of the β-catenin destruction complex, causing behaviour synonymous with a continuous Wnt signal [64, 91, 142, 171]. This leads to an aberrant build-up of β-catenin within the cell and the expression of target genes when we would otherwise expect basal expression.

A result of Wnt signalling in the crypt is that cell proliferation occurs predominantly towards the base, due to a low level of extracellular Wnt detected at the crypt orifice and a maximal level at the base [31, 91, 171, 175]. The resulting Wnt signalling gradient helps to maintain homeostasis in the crypt, with threshold levels of Wnt signalling linked to the propensity of cells to progress through the cell cycle [160]; thoroughly understanding the source of the Wnt gradient may help us better inform the growing field of intestinal organoid culturing to replicate intestinal crypts *ex vivo*.

Many *in silico* gastrointestinal modelling efforts have linked this positionally dependent and imposed active Wnt gradient directly to cell-cycle length [49, 93, 124, 173]. More recent experimental data suggest that the Wnt gradient results from the division-based spreading of Wnt from the base of the crypt, where there is an effective reservoir of Wnt, and consequent Wnt dilution throughout the crypt [56, 57]. In the case of the small intestine, this 'reservoir' can be considered akin to the region of differentiated, secretory Paneth cells that can transfer Wnt to neighbouring cells, although this diffusion has been shown to be limited to 1-2 neighbour cells [56, 57]. Additionally, a more recent study has suggested a possible source of Wnt proteins in FOXL1+ telocytes, which form a subepithelial plexus extending from the stomach to the colon, again localised towards the base of the crypt [153].

Aside from Wnt signalling, other mechanisms play important roles in crypt homeostasis, maintenance and dynamics in health and disease. It has been suggested that the Hippo signalling pathway, which is deregulated in multiple cancers [11, 26, 186], plays a role in the inhibition of cell proliferation due to contact-inhibition related mechanisms [92, 101], with signalling linked to the volume of a cell, and the overall stress applied to it [50, 92, 95]. The Hippo pathway negatively regulates Yes-associated protein (YAP) and the transcriptional co-activator PDZ-binding motif (TAZ); the activation of YAP and TAZ promotes cell proliferation and inhibits cell death [44, 188]. The Hippo pathway restricts the availability and the functionality of YAP in the nucleus by altering its distribution and protein levels [97, 187]. There are multiple studies that have established that the over-expression of YAP or its over-activation by Hippo pathway

mutations as able to counter the effects of cellular contact inhibition *in vitro* and organ size control *in vivo* to promote tissue overgrowth and cancer development [26, 188]. The Hippo signalling pathway, unlike other signalling cascades, does not appear to have its own dedicated extracellular peptides and receptors [89], relying on regulation by a network of upstream components and mechanisms, such as cell polarity complexes and adherens junctions [182]. It has been suggested that the Hippo pathway is regulated by cellular architecture and the mechanical properties of the environment, possibly serving as a sensor of tissue structure and mechanical tension [50, 137, 149].

Studies have shown that a cross-talk between the Hippo and Wnt signalling pathways plays a key role in mediating the proliferation of cells [92], by reducing nuclear β-catenin levels and, consequently, the expression of Wnt target genes that control cell-cycle progression [131]. Specifically, Hippo signalling reduces the nuclear localisation of β-catenin by increasing the concentration of phosphorylated YAP (YAP-P) in the cytoplasm; YAP-P binds to free cytoplasmic β-catenin causing it to be retained within the cytoplasm but without affecting the molecular stability of β-catenin [137]. In a Hippo 'off' state, the level of YAP-P remains stable within the cell, and as such the level of the YAP-P/β-catenin complex remains at a stable and minimal level in the cytoplasm. In a Hippo 'on' state, there is an increase in YAP-P, which binds to free β-catenin in the cytoplasm creating a YAP-P/β-catenin complex that is unable to localise to the nucleus of the cell whilst maintaining its membrane-bound activities [112]. This suggests a possible mechanism that links Wnt and Hippo signalling in contact inhibition, extending on a previously described molecular kinetic models of β-catenin within an individual cell.

In addition to the Wnt and Hippo signalling pathways, the Notch signalling pathway also plays a vital role in the homeostasis of the crypt, coordinating cell fate specification via interaction with the Wnt pathway [15, 64, 99, 174]. Crypt cells adopt a secretory (goblet, Paneth and enteroendocrine) or absorptive (enterocyte) phenotypes depending on the state of the Notch pathway, and its interplay with the Wnt pathway. As previously mentioned (Section 5.2.2) Paneth cells are secretory cells which undergo downward transit within the crypt, subsequently residing at the base of the crypt and provide a source of Wnt proteins. A recent study has suggested that the Notch target Hes1, and the protein Hath1 which is normally suppressed by Hes1, play an important role as they are affected by both the Notch and Wnt pathways [99]. A high expression of Hath1 (with corresponding suppression of Hes1) has been linked to secretory phenotypes, whilst a high expression of Hes1 has been seen in absorptive cells. The interplay between the Wnt and Notch pathway has the capability to determine the differentiative path of the crypt epithelial cells and thus aids the governing of crypt homeostasis.

5.2.4 Contact Inhibition

In cultured non-cancerous epithelial cells, the process by which cells transition from being freely proliferating and non-confluent (not in contact) to non-proliferating dense epithelial

monolayers is commonly referred to as contact inhibition (CI) [1, 5, 116]. In Puliafito *et al.* [5], CI within confluent cell cultures is defined as a "dramatic decrease of cell mobility and mitotic rate with increasing cell density as well as being the establishment of a stationary post-confluent state which is insensitive to nutrient renewal". Firstly this means that cell movement within the culture is reduced along with its cell-cycle progression, reducing the rate of proliferation. Secondly, it suggests that the reduction in the proliferation rate of the cells, as a result of the reduction in mitotic rate, causes cells to become insensitive to the level of nutrients in the environment. Additionally, solid stress has been shown to inhibit the growth of cell spheroids grown *in vitro*, regardless of cell differentiation state [82]. The mechanisms driving CI and the interplay of relevant signalling pathways involved in CI in the crypt are not thoroughly understood [187, 188]. Extensive work has been done to understand if and how signalling pathways and contact inhibition contribute to the mechanisms behind organ size control and cancer development [44, 188].

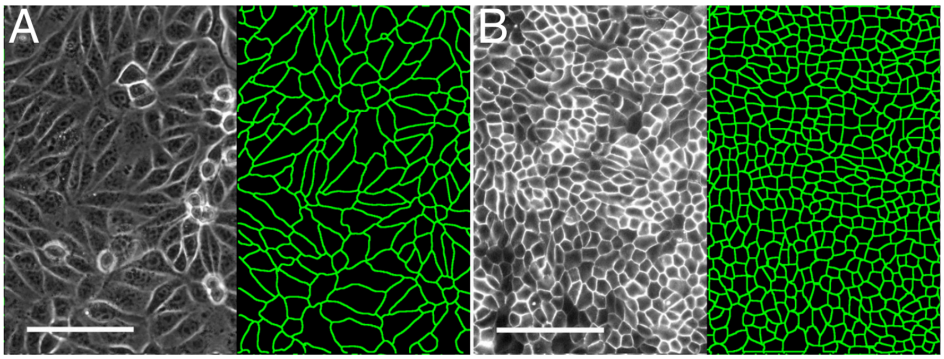


FIGURE 5.3. **Example of experimental imagery providing large-scale quantitative characterization of contact inhibition.** Image segmentation for MDCK (Madin-Darby canine kidney) cell cultures grown on PDMS. (A) Phase-contrast image of Ecad-GFP MDCK at low cell density. (B) Fluorescent image of Ecad-GFP MDCK at high cell density. Scale bar: 100μm. Reproduced from Puliafito *et al* 2012 [5].

Figure 5.3 provides an example of cells contact inhibited or not. The image in Figure 5.3 A show cells that are not undergoing CI, with their geometry being typically elongated, whereas Figure 5.3 B shows cells that are compressed and therefore are undergoing CI. These cells are more spherical and significantly smaller (1/6th) than the previous case.

It is widely accepted that CI requires the establishment of E-cadherin mediated cell-cell contacts [101, 182]; however, the nature of signalling pathways leading to this suppression remains unclear. One possible pathway, linking to this E-cadherin mediation, is the Wnt/β-catenin pathway. This makes sense as one of the key roles of β-catenin is in the mediation of adhesion at the cell boundary. Also, there is the possibility of a role involving the MAPK pathway, as this regulates the expression of cyclinD1 [58]. It is also suggested that the inhibitory effects

of crowding, causing the cells to become contact inhibited, is reversible. This means that if the stress on the cells is reduced, either by the removal of cells or by increasing the space they occupy in culture, then cells will readily proliferate again.

In Section 5.3 we describe the formation of a colonic crypt model, incorporating a gradient/threshold based representation of Wnt signalling within the crypt linked to a stochastic fixed-length cell-cycle model. To replicate the process of cellular CI we specify a volume dependent cell-cycle cessation method and investigate the effects of this initial simple representation of CI on the wild-type and mutant crypt dynamics.

5.3 Initial Dynamics in a Contact Inhibited Crypt

Having, in Section 5.2, described the biological systems governing the wild-type and dysplastic dynamics of small and large intestinal crypts, we extend on the modelling work reported in Chapter 4 by implementing and extending an agent-based model (ABM) of a large intestinal crypt, which incorporates mechanistic interactions between neighbouring cells and the mechanism of CI to govern the spatial dynamics of the crypt. As with the ABM in Chapter 4, we implemented our multiscale model in the Chaste simulation framework v3.2 [138].

5.3.1 Subcellular Signalling and Cell Cycle Interplay

As a number of previous modelling attempts have represented the subcellular Wnt dynamics in a gradient-dependent manner, we begin our investigation of intestinal crypts by implementing a previously described model [124], providing a baseline model to compare with when we extend it to incorporate CI. As described in Chapter 4, Section 4.3, we can define our ABM in terms of three specific modules, the subcellular/genetic model, the cell-cycle model and the mechanical model.

The subcellular/genetic representation of Wnt that we considered in this model is an external Wnt gradient [74], with a non-dimensional concentration W which decreases linearly from 1 as the base of the colonic crypt to 0 at the crypt ceiling. This Wnt gradient corresponds to the Wnt expression seen experimentally within a large intestinal crypt of a mouse [159]. Cells that reside towards the base of the crypt, in a high Wnt region, are assumed to produce the Wnt-dependent cell-cycle progression proteins, alluded to in Section 5.2.3, and therefore are able to progress through the cell cycle and subsequently divide. When a cell transits towards the top of the crypt and moves out of the region of high Wnt over a Wnt threshold W_{thr}, it is unable to produce Wnt-dependent cell-cycle progression proteins and therefore is considered differentiated. This threshold Wnt level (W_{thr}) was chosen such that the number of proliferative cells within the crypt matched experimental observations in Sunter *et al* [159]; thus we set $W_{thr} = 0.65$, resulting in cells proliferating within the lower 35% of the crypt as represented by the yellow region in Figure

5.4. If a cell is able to re-enter the proliferative region towards the base of the crypt then it can subsequently return to a proliferative state and continue to progress through its cell cycle.

Linking to the defined proliferative/differentiated regions of the crypt, the model that we considered is a fixed length cell-cycle model, in which a cell-cycle length is allocated to each cell at the point in time that it is created within the system [124]. As described in Chapter 4 Section 3.1.2, we considered the cell cycle in terms of its constituent phases, G_1, S, G_2 and M. In order to include an element of stochasticity into the cell-cycle process, and to prevent the synchronisation of cell division, we prescribed a G_1 length drawn from a normal distribution $N(\mu,1)$ where μ is fitted to experimental data [159]. From this data, the G_1 phase duration was sampled from a normal distribution $N(9.4,1)$h (truncated to ensure strictly positive cell-cycle times). The subsequent phases of the cell-cycle were $S = 7.4$h, $G_2 = 1.4$h and $M = 0.72$h. When a cell reaches the end of its cell cycle it divides, at which point new cell-cycle lengths are allocated according to the normal distribution. Cells that have left the Wnt high region of the crypt are considered to be perpetually in the G_0 phase of the cell cycle and never divide.

In considering the mechanical interactions within the crypt (i.e. cell-cell and cell-substrate) we applied an existing mechanical model [120], defined in Chapter 2 Section 2.3.1.3. As defined previously, the model describes cells as individual point centres connected to neighbouring cells (defined by performing a Delaunay triangulation) by linear over-damped springs. The individual cell morphology is subsequently defined by performing a Voronoi tessellation on the population of point centres. Unlike our agent-based model in Chapter 4, we are interested in how the physical composition of the cells, and their mechanistic states, affect the crypt epithelial cell population dynamics. We considered, in greater detail, this increased cellular mechanical requirement because we replicated the system as it exists *in vivo*, unlike in Chapter 4 where we were modelling a population culture *in vitro*.

5.3.2 Intestinal Crypt Takeover

Having defined our colonic crypt ABM we can now look at how it captures the wild-type and dysplastic dynamics observed experimentally [159]. We considered the process of mutant cell colonisation, or 'takeover', whereby at the end of the simulation the mutant cell has taken over the crypt replacing all non-mutant cells. Figure 5.4 shows snapshots providing an example of the temporal dynamics of takeover event which occurred over the course of a single simulation of our ABM of the crypt. At the beginning of the simulation, an initial single mutant cell was placed along the bottom row of the crypt (blue cell), with the proliferative and differentiated regions represented by the yellow and red cells respectively. By the end of the simulation, the crypt was entirely occupied by mutant cells. We were interested in how the probability of a mutant cell colonisation event changes with an increase in the advantage (represented by the mutant adhesion scaling term α) of the mutant cell.

We initially reproduced the results shown in [124]; this allowed us to set up the simulator

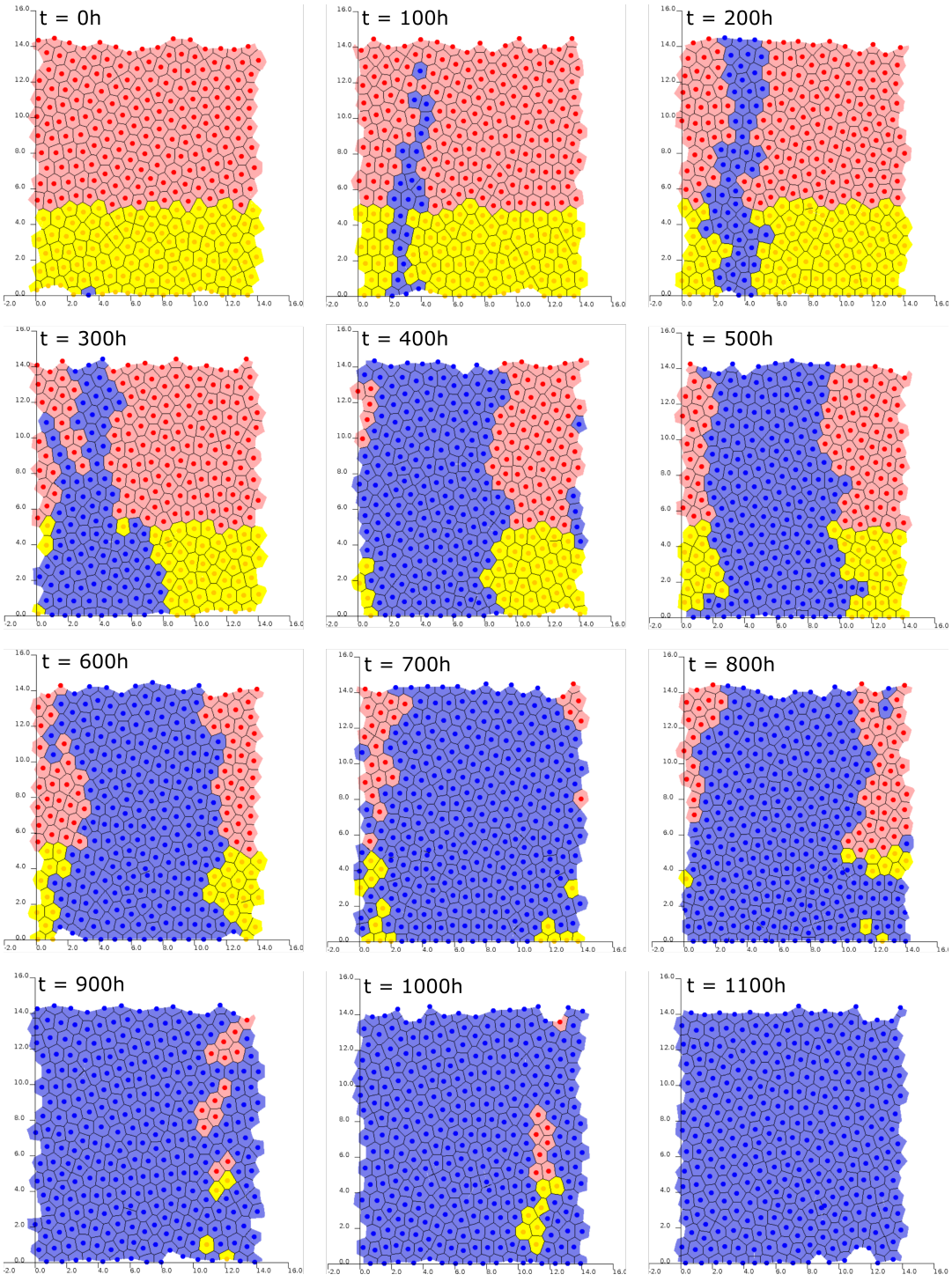


FIGURE 5.4. **Visualisation of a single example 'takeover' event in our agent-based model of the crypt.**. Snapshots taken from a single instance of mutant takeover in 100h intervals. The yellow cells are considered 'proliferative', the red cells are 'differentiated', and the blue cells are the 'mutant' cells. At the end of the simulation, the crypt is entirely occupied by the 'mutant' cells.

for later including our model of contact inhibition. Figure 5.5 shows the results of our replicated simulations, effectively looking at how the probability of crypt takeover increases with the increase in adhesion of the introduced mutant cell. These initial simulations were repeated 100 times for each value of α, with the mean of each simulation set represented by the black dots, and the corresponding shaded grey region representing the 95% confidence interval for a binomial distribution with a probability p, according to $p \pm z_{0.975}\sqrt{p(1-p)/n}$, where n is the number of simulations, and $z_{0.975}$ is the 97.5th percentile of a standard normal distribution.

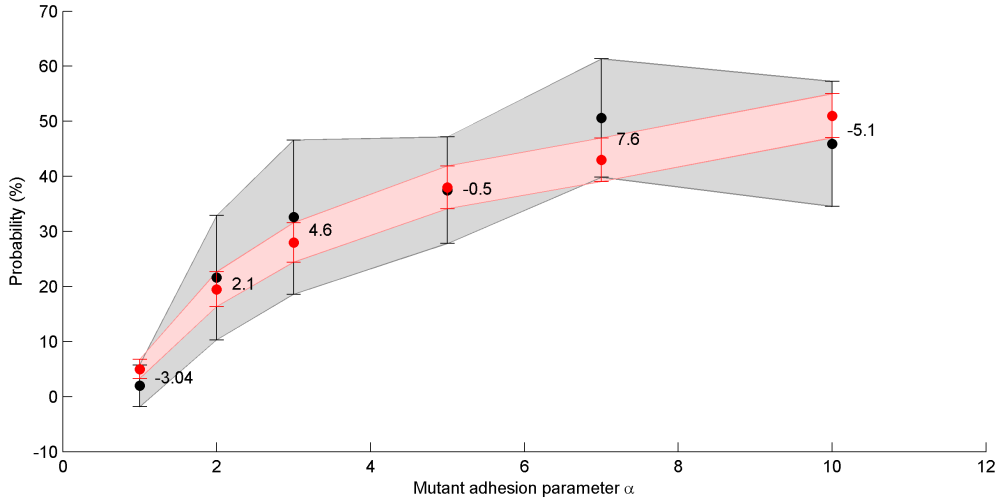


FIGURE 5.5. **Reproduced results from the model described in Mirams *et al* [124]**. The probability of mutant takeover of the crypt increases with an increase of mutant cell adhesion (α), shown as the red line and shaded error bars. The reproduced simulation results are shown as the black line and shaded error bar regions.

As the advantage of the mutant increased, with the increase of the cell-cell adhesion rate α, the probability of mutant takeover increased from approximately 5% when there is no advantage, to approximately 50% when $\alpha = 10$. The probability of mutant takeover increases with increasing α because the mutant cells are effectively subject to greater drag within the domain, causing them to occupy the crypt for longer and subsequently increasing the probability that they can out-compete the healthy cells. Our reproduced simulations are represented by the black shaded region, which matches the original simulations but with larger error regions due to a reduction in the number of simulations run.

5.3.3 Results of Contact Inhibition on Crypt Dynamics

Following this reproduction, we adapted the model to incorporate CI. In this model, we incorporated CI as a cell-cycle cessation method linked to the volume of the cell. When a cell volume drops below a threshold, cell-cycle progress is ceased until the cell volume has recovered to its equilibrium level. Figure 5.6 shows the dynamics from the simulations for both the original model and the introduced CI model, in terms of the changing cell counts of the three different cell-types (mutant, proliferative and differentiated). We have included the $\alpha = 1$ and $\alpha = 3$ cases, as beyond this value for the adhesion of the mutant cell the total cell-counts within the crypt become biologically unrealistic. Our values for CI were set as in a previous study [49], with healthy cells inhibited at 90% of their equilibrium volume (calculated by averaging the volume of dividing cells within the crypt), and mutant cells at 50% of their equilibrium volume. Figure 5.6(a) shows the washout dynamics when the only advantage for the mutant cell is this difference in CI threshold. It is clear from the cyan and red lines that the mutant cell is able to remain within the crypt for longer than in the original model, with an extension of approximately 200% (+400 hours) to the mean washout time for the mutant cell. Increasing this advantage by increasing the adhesion value to $\alpha = 3$ (Figure 5.6(b)), the washout time increased slightly, taking an additional 25% longer than the original model (+50 hours). Aside from this increase in the washout time, the dynamics for the other cell types remained the same, as expected.

Looking at the case of mutant cell takeover shown in Figure 5.6(c) and (d), we saw a greater difference between the dynamics in the original and extended CI models. In Figure 5.6(c), the time taken for the colonisation of the crypt by the mutant cell is vastly reduced when introducing CI into the model, with a reduction in monoclonal conversion time of approximately 73% when moving from the original to the CI crypt model. In addition to this, it appeared that in the original model there were oscillations in the cell-counts during the mutant takeover event (cyan line, Figure 5.6(c)) which were removed when considering the introduced CI model (red line, Figure 5.6). This shows a more direct takeover of the crypt by the mutant, due to the increased advantage of the mutant cell relative to the healthy cell. Subsequently considering the $\alpha = 3$ case (Figure5.6(d)), we see that the original model takeover time is reduced. When $\alpha = 1$, switching to our CI model reduces the takeover time by approximately 73% (Figure 5.6(c)), whereas switching to our CI model when $\alpha = 3$ the takeover time is reduced by approximately 50% (Figure 5.6(d)). Although the takeover event of the CI model is still significantly more rapid than with the original model, it also shows minimal signs of cell-count oscillations that are seen in the original model.

The results in Figure 5.6 indicate that the introduction of CI in the model vastly increased the advantage of the mutant cell and, subsequently, reduced the time taken for monoclonal conversion of the crypt. In addition, it appeared that the increased advantage gained by the mutant cells acted to reduce cell-count oscillations seen in the original model, which meant that the healthy cells are being inhibited and subsequently are unable to apply the required forces to the mutant population to push them out of the crypt.

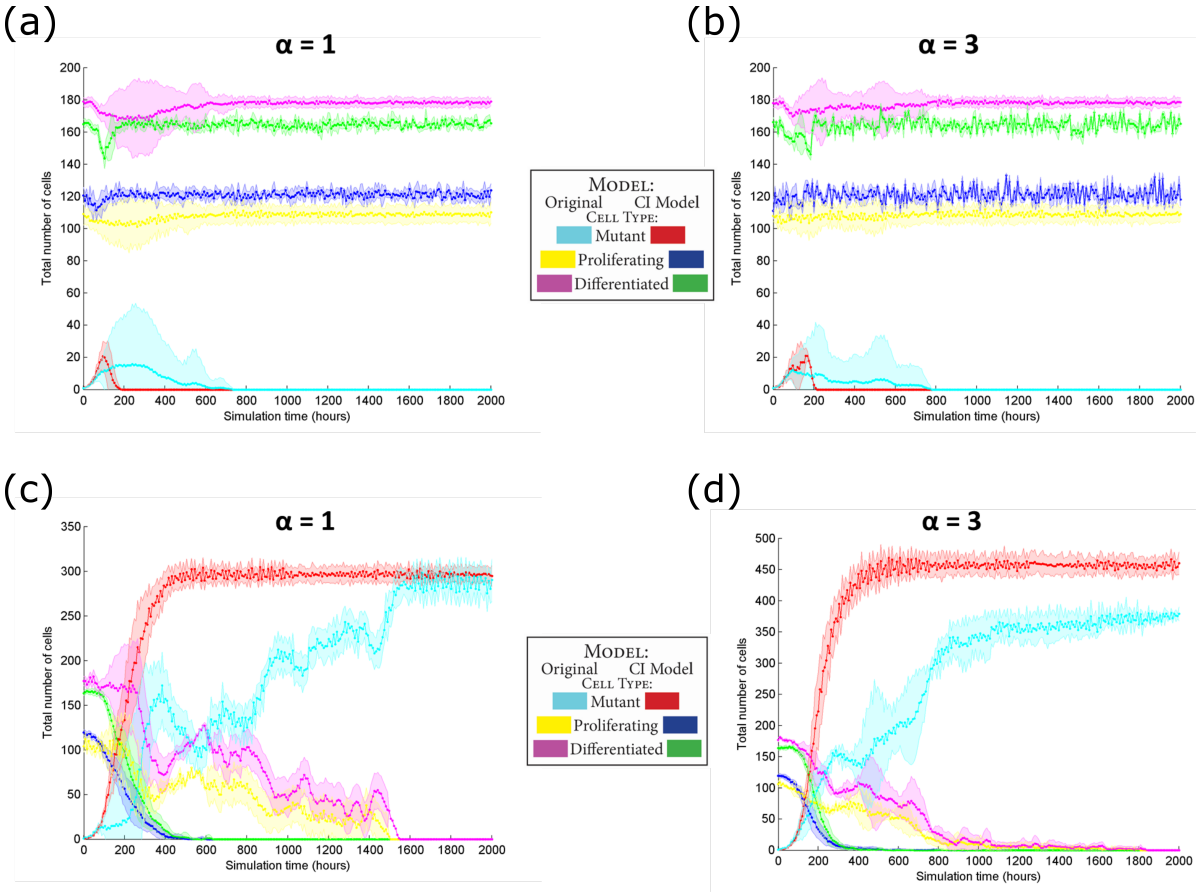


FIGURE 5.6. **Example takeover and washout dynamics in changing cell counts during the course of crypt simulation.** The cyan, yellow and purple lines represent the mean mutant, proliferating and differentiated cell counts, respectively, from the original [124] model. The red, blue and green lines represent the mean mutant, proliferating and differentiated cell counts, respectively, from the CI model. (a) Crypt washout dynamics for the three cell-types (mutant, proliferative and differentiated) over 2000h for the original model and the introduced CI model when $\alpha = 1$. (b) Crypt washout dynamics for three cell-types when $\alpha = 3$. (c) Crypt takeover dynamics for three cell-types when $\alpha = 1$. (d) Crypt takeover dynamics for three cell-types when $\alpha = 3$.

To further understand how these oscillations are affected by the introduction of CI into this model, we now look specifically at the dynamics of individual simulations; of note, the results in Figure 5.6 are the means of 150 simulations, and so the presence, and magnitude, of any oscillations could be hidden within this average.

In the simulations in Figure 5.6, we prescribed volume thresholds for CI of 90% equilibrium volume for healthy cells (H=0.9) and 50% equilibrium volume for mutant cells (M=0.5). In order to better understand the effect of CI on the transient dynamics of monoclonal conversion of

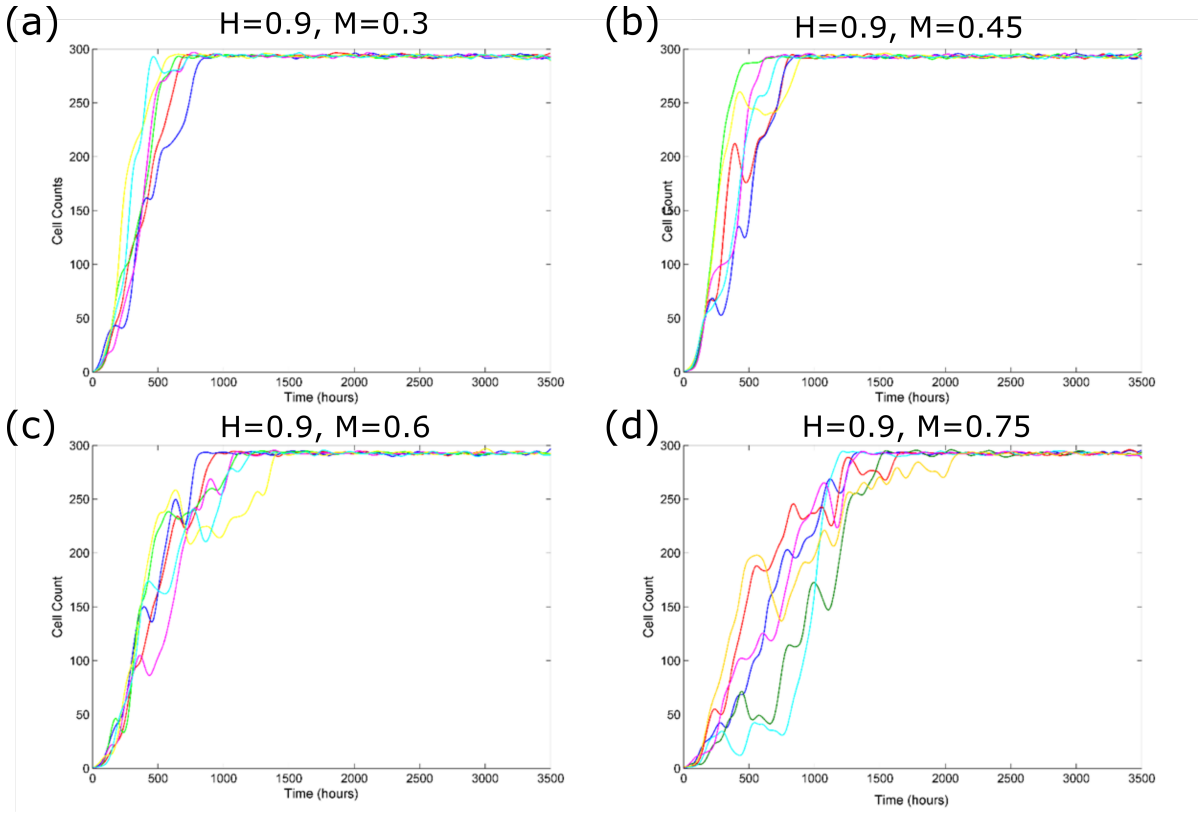


FIGURE 5.7. **Example simulation results showing the mutant cell-count oscillations within the extended CI model of the crypt.** (a) Cell-count oscillations from a sample 6 *in silico* crypt experiments, for a crypt with healthy CI at 90% of equilibrium volume and mutant CI at 30%. (b) Cell-count oscillations from a sample 6 *in silico* crypt experiments, for a crypt with healthy CI at 90% of equilibrium volume and mutant CI at 45%. (c) Cell-count oscillations from a sample 6 *in silico* crypt experiments, for a crypt with healthy CI at 90% of equilibrium volume and mutant CI at 60%. (d) Cell-count oscillations from a sample 6 *in silico* crypt experiments, for a crypt with healthy CI at 90% of equilibrium volume and mutant CI at 75%. H is the CI volume threshold for healthy cells, M is the CI volume threshold for mutant cells.

the crypt we varied these threshold volumes for CI for the mutant cells, and fixed the mutant adhesion parameter ($\alpha = 1$). Additionally, we considered individual simulations, as opposed to taking averages across simulations, to further understand whether the introduction of CI effectively removes the oscillations in cell count or whether the dynamics are hidden within the mean values across simulations. Figure 5.7 show examples of individual cell-count dynamics for mutant cells from our *in silico* crypt experiments. Figure 5.7(a) shows the least inhibited mutant cells, and as such cells with the greatest advantage over the healthy cells. In this case, there are minimal oscillations in the cell count and the event of monoclonal conversion is rapid (only

taking approximately 500h). Increasing this threshold to 45% equilibrium volume, as shown in Figure5.7(b), there are more pronounced oscillations and the time for monoclonal conversion takes an additional 200h compared to the lesser inhibited case (Figure 5.7). The individual cell counts shown in Figure 5.7(b) are effectively comparable with those used to calculate the mean counts in Figure 5.6(c). The oscillations in the cell-counts become more pronounced in Figures 5.7(c) and (d). In Figure 5.7(c) the time for monoclonal conversion is increased by approximately 500h from the original case (+100% longer), and there are pronounced oscillations during this takeover event, as the mutant cells are more readily inhibited and as such proliferation, and by extension their advantage over the healthy cells, is reduced. In figure 5.7(d) the takeover event takes significantly longer (+300% longer) and the oscillations in the cell-count are significant, with a maximal drop in mutant cell count of approximately 38% at one stage of a simulation. Considering the cell counts of mutant cells over during monoclonal conversion allows us to visualise the effect of increasing the advantage of our mutant cell compared to the healthy cells within the crypt, instead of switching a mutation 'on' we are gradually increasing the effectiveness of the mutation. Additionally, it has been suggested that this ribboning of cells within the crypt conform to one-dimensional neural drift dynamics, and so can be used to predict the number of stem cells residing within the crypt [6]. In a similar fashion, understanding the dynamics of these oscillations of same-lineage cells through the crypt in response to mutations targeting specific biological processes (such as those governing CI) can provide an insight into predicting the extent and severity of a crypt mutant.

5.3.4 Quantifying Crypt Motility Dynamics

In addition to considering the probability of mutant takeover and the changing interim dynamics seen in the form of crypt ribboning, we can further investigate the advantage gained by the mutant cell-type by introducing CI. An approach to this is to consider the changing cell motility, and we do this by tracking the movement of the cells over the course of the simulations. This allows us to quantify the speed at which they move throughout the crypt and in turn compare these motility speeds to those seen in the original ABM of the crypt without CI implemented.

Figure 5.8 shows the velocity differences between the original and the introduced CI threshold models for crypts where the introduced mutant has been washed from the crypt by the end of the simulation (i.e. between Figure 5.6(a) and Figure 5.6(b)). The velocities were averaged across the 2000 simulation hours, with 100 individual simulations completed. It is clear that as the threshold for healthy CI is increased, the velocity of the cells is reduced, with the minimum reduction occurring at the lowest threshold of CI (0.6) from the lower middle of the crypt (>25% crypt height) to the top of the crypt with a reduction of approximately 3% (Figures 5.8(b), (c) and (d)). The velocity difference is greatest, as expected, when the threshold for CI is increased to the maximal value (0.9) at the top of the crypt, where the cells would be expected to be moving fastest and therefore any reduction in proliferation has become more pronounced at this point (Figures

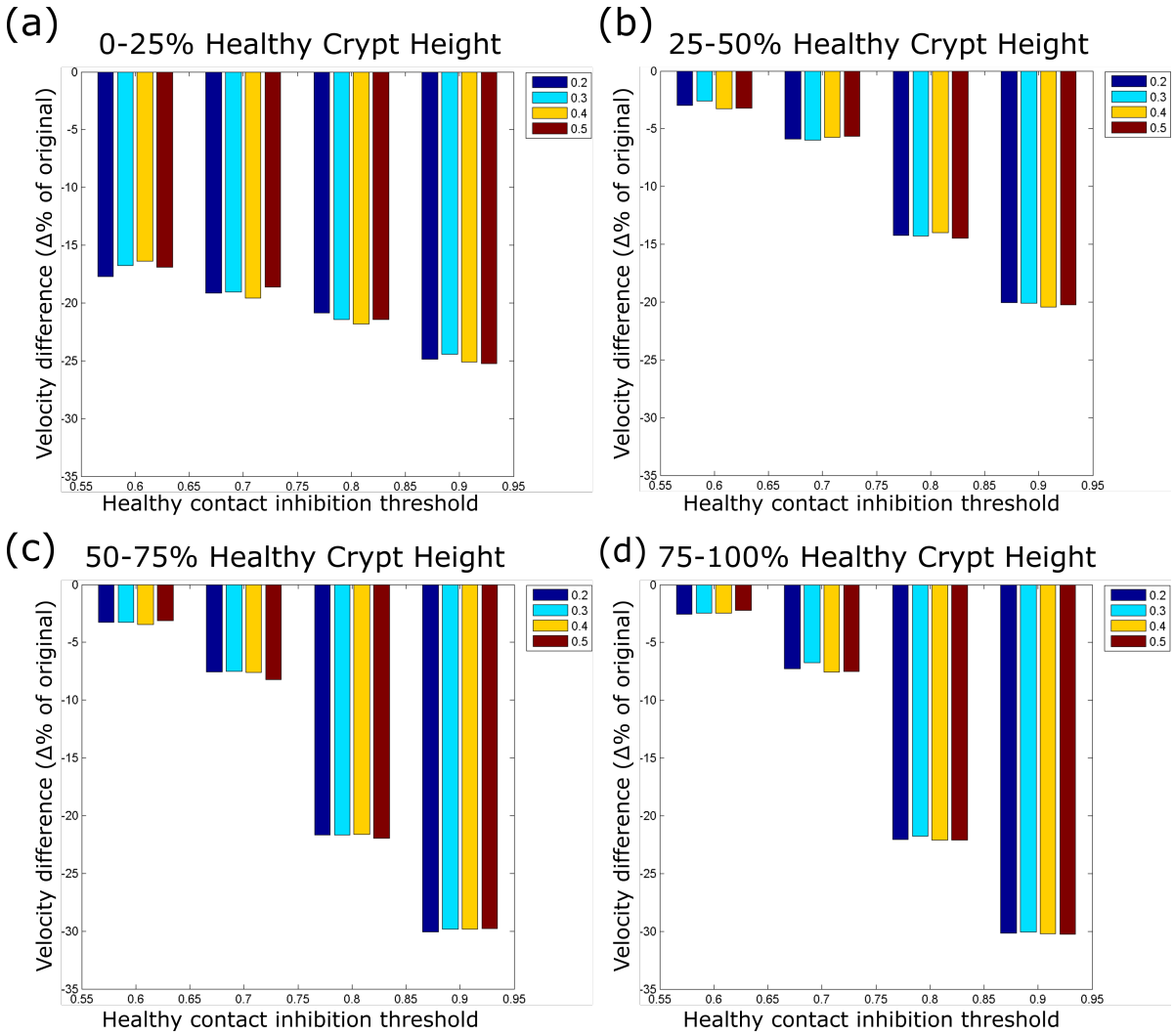


FIGURE 5.8. **Velocity differences in wild-type crypt cells at increasing heights in the crypt after introducing CI**. Velocity differences between healthy cells in the original crypt ABM [124] when introducing CI, for varying thresholds for healthy CI (indicated along the x-axis) and for mutant cell CI (blue, cyan, yellow and red bars). Velocities averaged within bottom 25% of the crypt (a), lower middle 25-50% of the crypt (b), upper middle 50-75% of the crypt (c), and top of the crypt 75-100% (d).

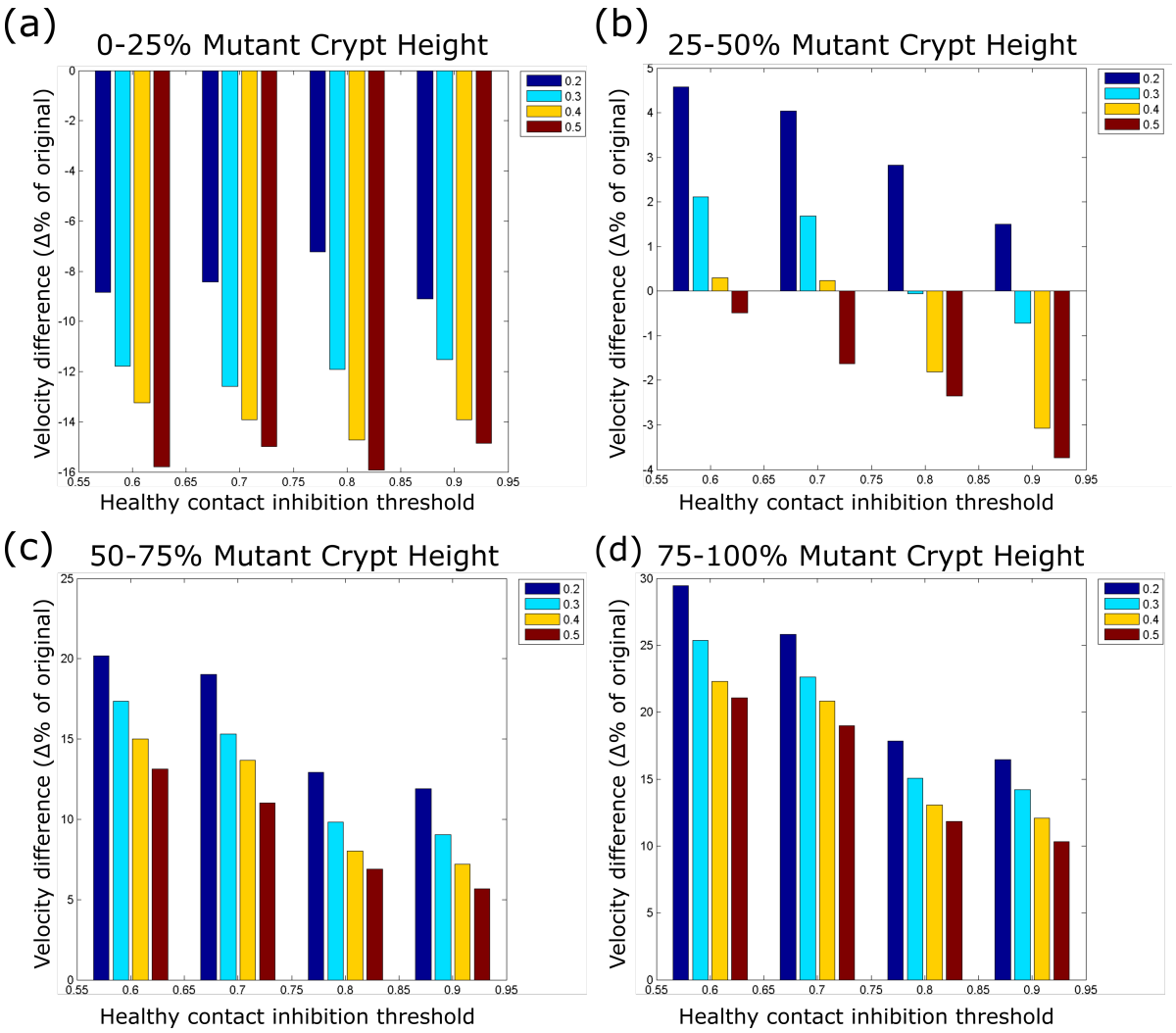


FIGURE 5.9. **Velocity differences in mutant cells at increasing heights in the crypt after introducing CI**. Velocity differences between mutant cells in the original crypt ABM [124] when introducing CI, for varying thresholds for healthy CI (indicated along the x-axis) and for mutant cell CI (blue, cyan, yellow and red bars). Velocities averaged within bottom 25% of the crypt (a), lower middle 25-50% of the crypt (b), upper middle 50-75% of the crypt (c), and top of the crypt 75-100% (d).

5.8(c) and (d)). At the top of the crypt (Figure 5.8(d)), the velocity of the cells is on average 30% slower than in the original model, which in turn means that the forces applied to mutant cells to remove them from the crypt are reduced. This reduction in velocity can be attributed to the reduction in proliferation due to the introduction of CI, and as such the dividing cells apply a reduced force to their neighbours which propagates as cells move upwards in the crypt. As expected, in the healthy (i.e. wild-type) crypt, the difference of increasing mutant cell CI threshold volume (moving from blue to red bars) does not affect the velocity of the cells in the crypt, as they only reside within the crypt for at most approximately 800h, as shown in Figure 5.6(b). We then considered mutant takeover dynamics upon introducing CI. Figure 5.9 shows the resultant velocity differences seen in a crypt taken over by mutant cells (i.e comparing velocity differences between simulations in Figure 5.6(c) and Figure 5.6(d)). In these simulations, we are comparing the crypts with introduced adhesive mutants ($\alpha = 3$) with increasing levels of mutant and healthy cell CI. The greater the difference between the CI thresholds of the healthy and mutant cell types the greater the advantage, similar to increasing the value of α. As compared to Figure 5.8, increasing the threshold volume at which healthy cells are inhibited has a lesser pronounced effect on the velocity of the cells within the crypt, and increasing the threshold volume of the mutant cells (from 0.2 to 0.5 x equilibrium volume) causes a decrease in the velocity of the cells (all panels Figure 5.9). This is to be expected, as in this mutant takeover case both healthy and mutant cells reside within the crypt for a not-insignificant time over the course of the simulation (as shown by the persistent presence of both healthy and mutant cells in Figure 5.6 (c) and (d)). Within the bottom 25% of the mutant crypt, the major difference to the cell velocity comes from increasing the mutant CI threshold (Figure 5.9(a)). As you move further towards the top of the crypt, both the mutant and healthy cell thresholds affect the cell velocities (Figures 5.9(b)-(d)).

It is clear from these results that the inclusion of CI in the model is able to alter both the wild-type (Figure 5.8) and mutant (Figure 5.9) cell motility within the crypt, slowing down the movement of cells through the crypt and subsequently causing cells to reside within the crypt for a longer period of time. This prolonged stay within the crypt can be seen as a possible cause for an increased rate of mutant monoclonal conversion, as the mutant cell is able to reside within the crypt and proliferate for longer allowing for a greater probability of taking the crypt over. Additionally, the reduction in motility of the healthy cells also reduces the forces being applied to the cells within the crypt.

5.4 Subcellular Signalling Crosstalk as a Mechanism in Contact Inhibition

Although the results of the model described in Section 5.3 provide an insight into the dynamics of crypt dysplasia, showing that an increase in cell-cell adhesion combined with CI causes a vast increase in the rate of monoclonal conversion and alters the spatial-temporal behaviour of the

crypt by varying the magnitude of cell fluctuations (cell-count oscillation in Figure 5.7), the model makes significant simplifications to the underlying cell-biology. In order to consider the process of CI in further detail, we considered a more complex model that accounts for the mechanisms by which signalling pathway cross-talk results in CI and subsequently how this affects cells across the whole spatial domain of the crypt. In the following section, we look at a crypt model that incorporates subcellular signalling and extend this to incorporate a mechanism for CI driven by Hippo signalling and the cross-talk with the Wnt signalling pathway.

Specifically, we investigate the role of Hippo signalling, and its possible cross-talk with the Wnt signalling pathway, to mediate these biological processes. As with the previous ABM, we split this up into three distinct modules, extending the complexity of the subcellular and the cell-cycle model of the cell and, finally, investigating their coupling with spatial effects.

5.4.1 Subcellular Wnt/Hippo Signalling Model

As described briefly in Chapter 2 Section 2.3, a number of previous models of crypt dynamics have focussed on the role of Wnt signalling, including its role in monoclonal conversion, the process by which a resultant crypt is populated in its entirety by the progeny of a single ancestor cell [124], as well as its effect on subcellular β-catenin kinetics and links to the cell cycle [173]. The main assumption about Wnt in all these models is that Wnt acts on cells as an imposed external gradient [91]. Earlier models assumed that Wnt affects the cell cycle throughout its lifetime, while more recent work suggests that the cell-cycle length is dependent on the Wnt signal available immediately after mitosis.

In order to elucidate the effects of different Wnt models on crypt dynamics we considered two possible Wnt models: the first assumes that Wnt exists as an externally-imposed gradient, with variants as described above, where the Wnt level of the cell (on which the cell cycle depends) either changes throughout its lifetime or is fixed at the point of each cell division. The second model considers Wnt to be held internally within each individual cell, instead of being prescribed externally, and hence allows us to test the hypothesis of an emergent Wnt gradient in the crypt. In this second model, Wnt is divided between the daughter cells upon cell division, with a Wnt source 'reservoir' region at the bottom of the crypt in which cells can receive Wnt; as mentioned previously, in the case of the small intestine this is akin to the region of Paneth cells that can transfer Wnt to neighbouring cells, diffusion limited to 1-2 cell distances [56, 57].

For our *in silico* experiments these models are defined as follows: the first (labelled M_E) includes Wnt as an externally-imposed gradient, and has two variants: static (labelled M_{E1}), in which each cell reads its Wnt signal at birth, and dynamic (labelled M_{E2}), in which each cell continuously updates its Wnt level, as in previous models [124, 173]. In the second Wnt model (labelled M_I), each cell contains an independent level of Wnt, split at each cell division event between the two daughter cells.

To link these possible Wnt descriptions to our individual cells, and subsequently to cell-cycle

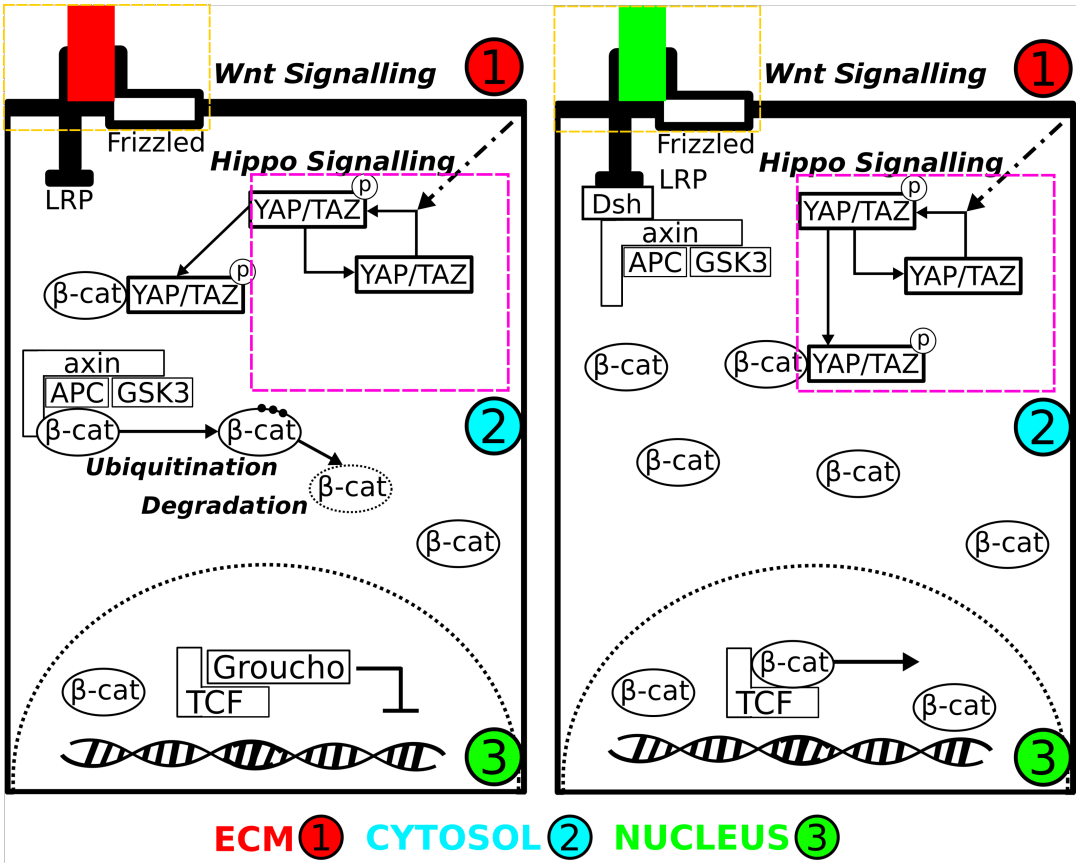


FIGURE 5.10. **Wnt/Hippo signalling network schematic.** Schematic showing β-catenin localisation, dependent on Wnt (yellow box) and Hippo (purple box) signalling. β-catenin exists either in the cytosol, at the extracellular matrix (cell surface) or in the nucleus, with its localisation directly affecting cell-cycle progression. In a Wnt-Off state (left panel) β-catenin is degraded by a complex comprised of APC, Axin and GSK3, preventing β-catenin localisation. In a Wnt-On state (right panel), the destruction complex is disrupted by Disheveled (Dsh) at the extracellular matrix, resulting in increased β-catenin levels and its nuclear localisation. This causes the displacement of Groucho in the nucleus and transcription of Wnt target genes linked to proliferation and cell-cycle progression. Hippo signalling acts to cause the phosphorylation of YAP/TAZ within the cytosol of the cell. Phosphorylated YAP/TAZ binds to cytosolic β-catenin, preventing nuclear accumulation, in turn preventing transcription of Wnt target genes and cell-cycle progression.

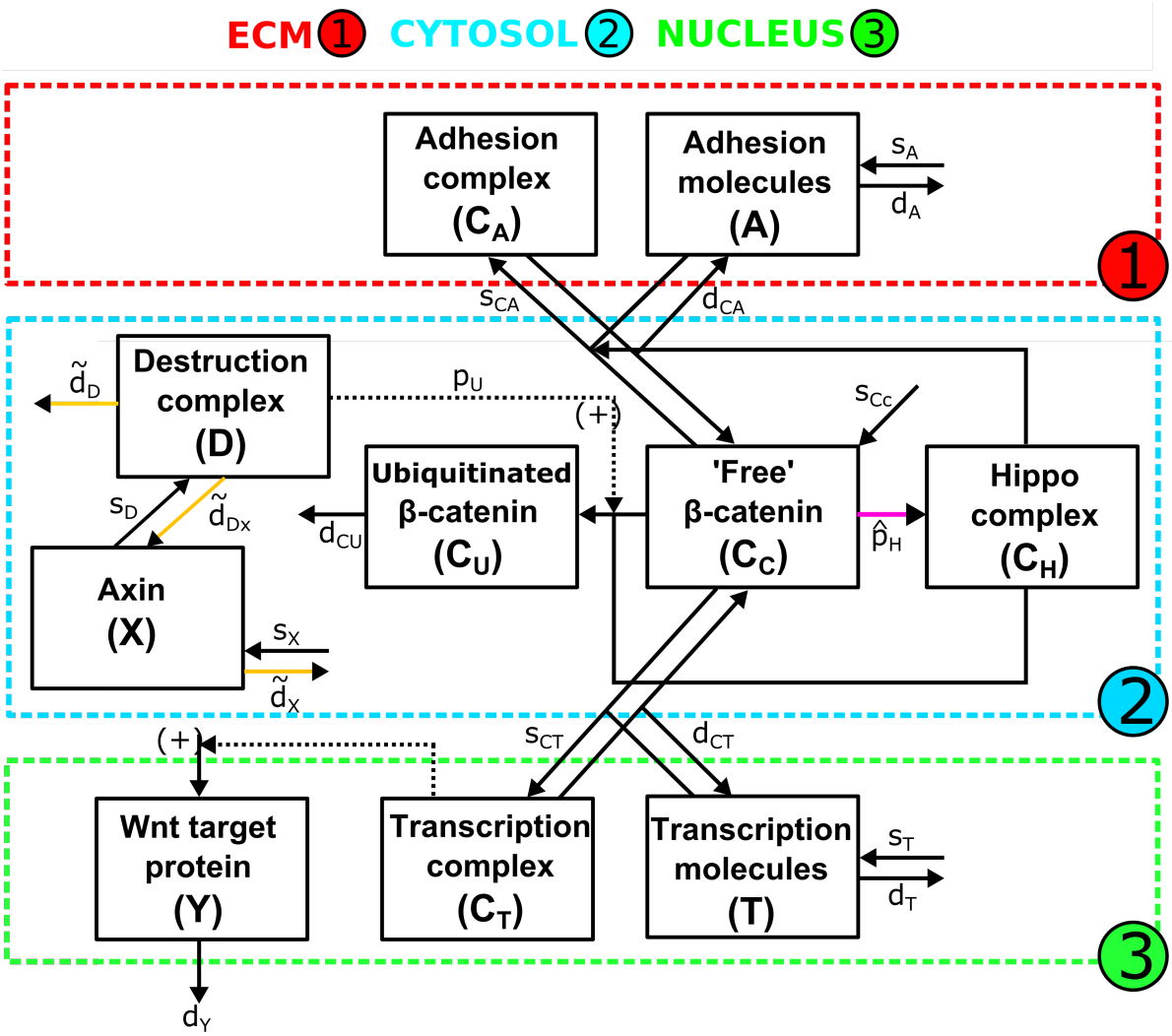


FIGURE 5.11. **Network describing Wnt/Hippo cross-talk within crypt epithelial cells.** Network diagram resulting from the schematic in Figure 5.10, describing the kinetics of β-catenin within the cell. C_A, C_C, C_T, and C_U are the levels of adhesive-linked β-catenin at the cell surface, cytosolic β-catenin, transcriptional 'nuclear' β-catenin, and β-catenin marked for degradation, respectively. A, T, and D denote the level of molecular species forming complexes with β-catenin at the cell surface (forming adhesive complexes at the adherens junction), within the nucleus, and within the destruction complex, respectively. X and Y denote the levels of Axin and transcribed Wnt target proteins, respectively. C_H denotes the level of β-catenin/YAP complex formed due to Hippo signalling in the cell. Parameter values and definitions in Table 5.1. Rates which depend on activity of signalling pathways are indicated by the coloured arrows (yellow for Wnt dependence, purple for Hippo dependence).

dynamics, we adapted a previously published [172] ordinary differential equation (ODE)-based kinetic model of Wnt-effector β-catenin. This allows us to capture the combined effects of Wnt/β-catenin and, additionally, it provides us with a base model to further include the effects of Hippo signalling on intracellular dynamics through localisation of β-catenin. Although alternative models of Wnt signalling exist [107, 113, 123, 135, 148], the van Leeuwen formalism was chosen as it incorporates important mechanistic features of the canonical Wnt pathway (including sequestration of β-catenin by the destruction complex, and activation/inactivation of the destruction complex) and couples β-catenin localisation with cell-cycle progression [160]. Specifically, when β-catenin localises at the cell membrane, it regulates the formation of E-cadherin-dependent cell-cell contacts, connecting adherens junction proteins to the actin cytoskeleton [172]. We investigated the role of Wnt in the wild-type and dysplastic dynamics of the crypt, and the interplay with Hippo. Figure 5.10 shows schematically the interplay between Wnt and Hippo signalling that forms the basis for the kinetic diagram (5.11), representing relations among relevant variables, namely: free axin (X), adhesive molecules (A), and transcriptional molecules (T), the active destruction complex (D), the Wnt target protein (Y), four molecular forms of (C_i, $i = A, C, T$ and U, corresponding to β-catenin contained in the adhesive junction, cytoplasm, transcriptional complexes, and marked for ubiquitination/degradation, respectively), and a β-catenin/Hippo complex (C_H). The symbols carrying tildes ($\tilde{d}_D$, $\tilde{d}_{DX}$ and $\tilde{d}_X$) are those that may vary in response to Wnt signalling. It was assumed that these parameters are linear functions of Wnt signal (W), which varies from $W = 0$ to $W = 1$ within the prescribed Wnt region (40%/100%). They are described as follows: $\tilde{d}_X(W) = d_X + \epsilon_X W$, $\tilde{d}_{DX}(W) = d_{DX} + \epsilon_{DX} W$ and $\tilde{d}_D(W) = d_D + \epsilon_D W$

APC-mediated degradation of β-catenin

$$\frac{d[D]}{dt} = s_D[X] - (\tilde{d}_D + \tilde{d}_{DX})[D] \tag{5.1}$$

$$\frac{d[X]}{dt} = s_X - s_D[X] - \tilde{d}_D[X] + \tilde{d}_{DX}[D] \tag{5.2}$$

$$\frac{d[C_U]}{dt} = \frac{p_U[D][C_C]}{[C_C] + K_D} + \frac{p_U[D][C_H]}{[C_H] + K_D} - d_U[C_U] \tag{5.3}$$

Cytoplasmic β-catenin

$$\frac{d[C_C]}{dt} = s_C + d_{CA}[C_A] + d_{CT}[C_T] - (s_{CA}[A] + s_{CT}[T] + d_C)[C_C] - \frac{p_U[D][C_C]}{C_U + K_H} - \frac{\hat{p}_H[C_C]}{[C_C] + K_H} \tag{5.4}$$

$$\frac{d[C_H]}{dt} = \frac{\hat{p}_H[C_C]}{[C_C] + K_H} - s_{CA}[C_H][A] - \frac{p_U[D][C_H]}{[C_H] + K_D} \tag{5.5}$$

E-cadherin-mediated cell-cell adhesion

$$\frac{d[A]}{dt} = s_A + d_{CA}[C_{CA}] - (s_{CA}[C_C] + d_A + s_{CA}[C_H])[A] \tag{5.6}$$

$$\frac{d[C_A]}{dt} = s_{CA}[C_C][A] - d_{CA}[C_A] + s_{CA}[C_H][A] \tag{5.7}$$

Transcription of Wnt-target genes

$$\frac{d[T]}{dt} = s_T + d_{CT}[C_T] - (s_C[C_C] + d_T)[T] \tag{5.8}$$

$$\frac{d[C_T]}{dt} = s_{CT}[C_C][T] - d_{CT}[C_T] \tag{5.9}$$

The temporal dynamics and interactions of these variables within a single cell are described as a system of ordinary differential equations (ODEs, Equations 5.1 – 5.9). They include adaptations to include effects of the Hippo signalling via the phosphorylated YAP/β-catenin complex as well as contact inhibition effects, described in Section 5.4.2. Key to the model is β-catenin, with its location within the cell controlling the transcription of Wnt target genes linked to proliferation, as well as the cell adhesive characteristics. Figure 5.11 shows schematically how β-catenin is able to localise throughout the cell. In the van Leeuwen et al (2007) [172] model, cytoplasmic can either: (a) be sequestered by molecules linked to Wnt gene transcription to form transcriptional complexes, (b) form adhesive complexes at adherens junctions, or (c) undergo Wnt-mediated degradation. We extended the model to include the effect on β-catenin of Hippo signalling in the cell, having linked Hippo and Wnt signalling. Specifically, as postulated by Imajo et al. (2012) [92] , we modelled the Hippo signalling-governed concentration of cytoplasmic phosphorylated YAP (YAP_P) to cause a reduction in nuclear β-catenin. This is in line with experiments showing that Hippo signalling causes increased phosphorylation of YAP, combined with the formation of YAP-P/β-catenin complexes within the cell [11, 92, 187]. It has also been suggested that the β-catenin/YAP_P complex can localise to the membrane [12, 92] The β-catenin/YAP_P complex is incorporated in the ODE system by the inclusion of Equation 5.5. We described the formation of this complex as being governed by Hill kinetics, with the rate of formation of the complex controlled by a Hippo-dependent signal. We assumed that the mechanism of CI is a reduction of cell-volume, as this should correlate with an increase in surface-stress due to cellular crowding in a cell layer, as well as an increase in cell density within a layer. We therefore prescribed the rate $\hat{p}_H$ (nuclear sequestration rate of β-catenin [92]) to be volume-dependent, which, in turn, describes how volume-dependent Hippo signalling decreases the nuclear localisation of β-catenin (and, as such, increases its cytoplasmic presence, without altering the total concentration of

β-catenin within the cell) and, consequently, cell proliferation dynamics. Table 5.1 shows the parameters used in this network description of β-catenin cellular kinetics, including a description of each of the parameters.

5.4.1.1 Fitting and Sensitivity Analysis of Extended Subcellular Network Model

To explore the effect of Hippo signalling on cell-cycle length, we performed a parameter sensitivity analysis, varying the YAP-P/β-catenin complex binding rate (p_H) and dissociation constant (K_H) (Equation 5.5); these parameters govern the effectiveness of the inhibition of nuclear β-catenin localisation and, consequently, contact inhibition. The results are shown in Figure 5.12(a). They suggest that the saturation coefficient K_H has a small effect on the cell-cycle length but alters the effective range over which the binding rate p_H has to vary, with the binding rate then having the largest effect on lengthening and eventually stopping cell-cycle progression. We, therefore, assigned a value to the saturation coefficient K_H which, when combined with Wnt signalling effects on β-catenin localisation, did not cause premature cessation of the cell cycle. The parameter p_H is varied within a given range (i.e. takes specific values as cell volume is not directly modelled), while, in the agent-based model, $\hat{p}_H$ is dependent on cell volume (i.e $\hat{p}_H = p_H(V)$. Single-cell simulation results, suggest that the saturation coefficient (K_H) has some some effect on cell-cycle duration but, more significantly, alters the effective range over which the binding rate (p_H) controls cell-cycle duration, with changes in the binding rate having the largest effect on delaying and eventually stopping proliferation.

We then investigated how the cross-talk between the Hippo and Wnt pathways affects the length of the cell cycle (Figure 5.12(b)). Fixing the saturation term K_H (set to be $20nM$, to allow for an appropriately sized range over which to vary the binding rate), a concurrent increase in the YAP_P/ β-catenin complex binding rate pH, and decrease in Wnt signalling, causes a lengthening and eventual cessation of the cell cycle (with the length of the G_1 phase of the cell-cycle determined by the time taken for the level of E2F1 to surpass the threshold $E2F1 > 1$), suggesting that cells with lower Wnt concentrations are more susceptible to being inhibited by Hippo signalling as there is less free β-catenin localised to the cytoplasm.

The results from this single cell model analysis show that the reduction in cytoplasmic caused by Hippo signalling is able to prevent the progression through the cell cycle, thus effectively inhibiting proliferation.

5.4.2 Crypt Modelling Setup

Whilst describing the cell cycle in terms of these ODEs is clearly more complex than prescribing a stochastic-deterministic cell cycle model, however, in van Leeuwen *et al.* this model is used to produce biologically relevant results [173]. In further models that investigate intestinal crypt dynamics [49, 124] they forgo this more complicated model and revert to the simpler prescribed cell cycle case. This is, therefore, an interesting question to address, whether this

TABLE 5.1. **Parameters for the kinetic Wnt/Hippo signalling model.** Model parameters for system expressed in Equations 5.1 – 5.9 describing the change of β-catenin localisation within the cell. s_C is the rate of production of β-catenin within the cytosol, s_{CA} is the formation rate of adhesive β-catenin complexes, s_{CT} is the formation rate of transcriptional β-catenin, s_D is the rate of formation of the destruction complex, s_X is the basal rate of axin production, s_T is the basal rate of transcriptional molecule production, s_A is the basal rate of adhesive molecule formation and s_Y is the maximal rate of transcription of Wnt target genes. Additional parameters are as follows, d_{CA} is the disassociation rate of adhesive complexes, d_{CT} is the disassociation rate of transcriptional complexes, d_T and d_A are the destruction rates of transcriptional and adhesive molecules respectively, $\tilde{d}_X$ and $\tilde{d}_{DX}$ are the Wnt-dependent degradation rate for axin and disassociation rate for the destruction complex respectively, $\tilde{d_D}$ is the Wnt-dependent degradation rate for the destruction complex, d_Y refers to the transcription of Wnt target genes. P_U is the binding rate for the destruction complex and cytosolic β-catenin, K_D is the dissociation constant for the destruction complex, p_H is the binding rate for phosphorylated YAP (YAP_P) and β-catenin, K_H is the dissociation rate for the YAP_P-β-catenin complex. p_{Hmax} is the Hippo (hence volume) dependent association rate for YAP-P and β-catenin. Δ describes the volume reduction required to cause complete cell-cycle cessation.

$s_A = 20nM/h$	$d_A = 2h^{-1}$	$\tilde{d}_X = 100h^{-1}$	$\epsilon_{DX} = 5h^{-1}$
$s_{CA} = 250(nMh)^{-1}$	$d_{CA} = 350h^{-1}$	$d_Y = 1h^{-1}$	$\epsilon_D = 5h^{-1}$
$s_C = 25nM/h$	$d_C = 1h^{-1}$	$K_D = 5nM$	$\Delta = 0.2$
$s_{CT} = 30(nMh)^{-1}$	$d_{CT} = 750h^{-1}$	$K_T = 50nM$	
$s_D = 100h^{-1}$	$\tilde{d}_D = 5h^{-1}$	$p_U = 100h^{-1}$	
$s_T = 10nM/h$	$\tilde{d}_{DX} = 5h^{-1}$	$K_H = 20nM$	
$s_X = 10nM/h$	$d_T = 0.4h^{-1}$	$p_{Hmax} = 26000h^{-1}$	
$s_Y = 10h^{-1}$	$d_U = 50h^{-1}$	$\epsilon_X = 200h^{-1}$	

added complexity will yield suitably more accurate results when adapted to incorporate additional signalling pathways and whether the trade-off added parameter complexity is worth it. In our initial investigation in Section 5.3 we will consider this simple cell-cycle model, extended to include cellular contact inhibition.

We coupled the cell-cycle model to the Wnt/Hippo subcellular network model to account for contact inhibition effects: the YAP_P/β-catenin complex is able to cease cell-cycle progression by decreasing the level of free cytoplasmic β-catenin. Specifically, we coupled the cell-cycle model to the level of transcriptional complexes, C_T (Figure 5.11), as described in Swat *et al*. The model coupled cell proliferation to levels of cyclin D (described in the model by the variable Y in Figure 5.11), which is a known Wnt target that promotes the transition from the G_1 to S phases of the cell cycle. The greater Wnt signal at the base of the crypt causes the cell to produce cyclin D and

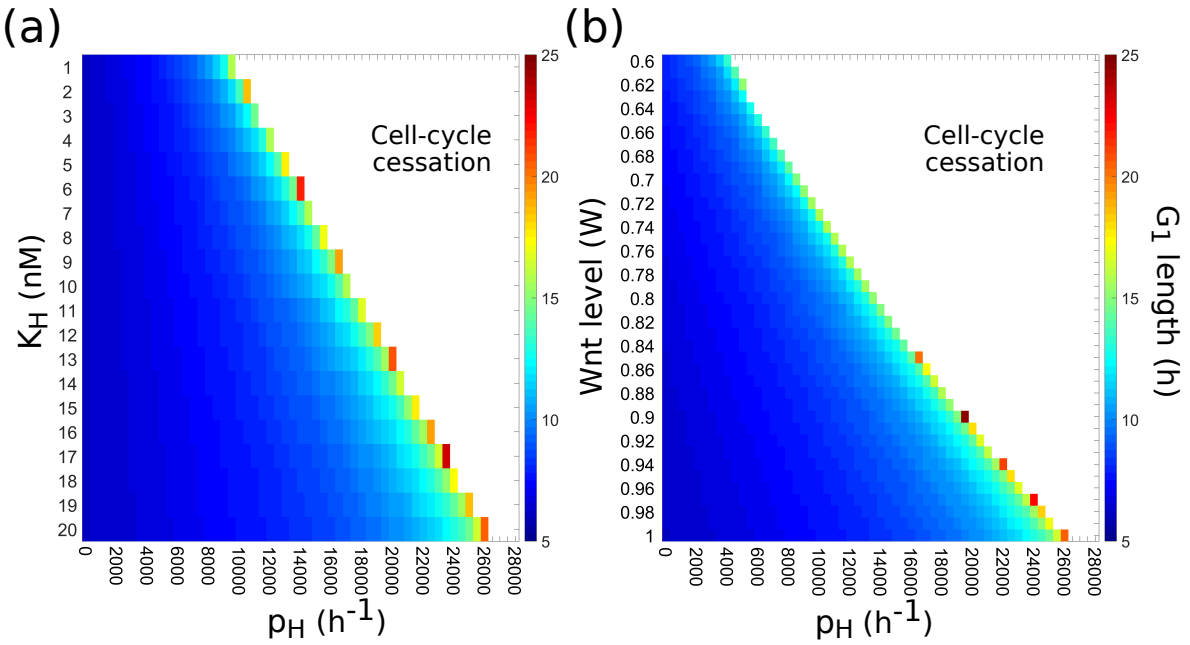


FIGURE 5.12. **Single cell cell-cycle response to Hippo signalling.** (a) Single-cell sensitivity analysis of the Hippo signalling module; parameters p_H and K_H represent the binding and dissociation rates of phosphorylated YAP/β-catenin and its complex, respectively. The coloured region shows the length of the cell cycle during progression, and the white region represents cell-cycle cessation. Increasing the value of K_H affects the cell-cycle length less aggressively than p_H, it instead increases the range over which p_H can exist whilst continuing cell-cycle progression. (b) Single-cell sensitivity analysis of the Hippo-dependent cell cycle as a function of (static input) Wnt level and p_H. The effective range for p_H shortens with reduced Wnt signal, suggesting that this Hippo model is more aggressive at lower levels of Wnt, towards the end of the proliferative zone of the crypt.

thus progress through the cell cycle more rapidly, ultimately creating an increasing age gradient over the crypt, and the appearance of a proliferative base of the crypt, as seen experimentally [159]. This mechanism causes a competition between the two pathways: an increase in Wnt shortens the length of the cell cycle while increasing Hippo causes cell-cycle lengthening.

Although our single-cell analysis of the effect of this Hippo signalling on cell-cycle length (shown in Figure 5.12) provides some information as to the interplay of these two pathways, we can better understand how this develops within the crypt system by implementing it within an ABM. We made the rate p_H volume dependent ($\hat{p}_H$), varying depending on the cell volume (V), according to:

$$\hat{p}_H = p_H(V) = \begin{cases} 0, & V \geq V_{thr} \\ -\frac{p_{Hmax}}{\Delta V_{thr}} V + \frac{p_{Hmax}}{\Delta}, & \Delta V_{thr} \leq V < V_{thr} \\ p_{Hmax}, & V < \Delta V_{thr} \end{cases} \tag{5.10}$$

where V_{thr} is the threshold volume for the onset of contact inhibition, p_{HMax} is the maximal rate of Hippo dependent β-catenin sequestration (set using the single cell model as shown in Figure 5.12), V_{thr} is the volume threshold at which Hippo is active, and Δ describes the volume reduction required to cause complete cell-cycle cessation. The value for Δ is fitted to allow cell-cycle length elongation before cessation, with $\Delta = 0$ effectively causing contact inhibition to occur instantaneously when the cell volume drops below the threshold V_{thr}. Including $\Delta > 0$ maintains the proliferative region towards the crypt base. The process of cell-cycle cessation due to volume reduction has been previously implemented as a switch [49], stopping cell-cycle progression below a set volume, while our implementation results in smoother transitions.

The agent-based simulations (shown in Figures 5.14, 5.15 and 5.16) were conducted with a time step of 0.01h; each experiment lasts for 1000 (simulated) hours, and was repeated 150 times to provide sufficient data for statistical analysis.

As with the simple crypt model described in Section 5.3 we maintained the same description of the mechanics of the cells within our crypt domain [120], as the cell-centred approach has been shown to yield biologically relevant results in previous modelling attempts [48, 60, 124, 173]. Figure 5.13 provides a visualisation of our crypt ABM, describing the 3D to 2D domain simplification and the respective periodic boundaries, as well as indicating the differences between the Wnt gradients that we considered. Additionally, it presents the different sizes of the proliferative regions in the crypt with the two different Wnt gradient sizes.

In the following section, we investigate the effect of our two different Wnt models, along with the inclusion of our Hippo signalling, on the wild-type crypt homeostasis.

5.4.3 The Effects of Hippo and Wnt Signalling on Wild-Type Crypt Homeostasis

5.4.3.1 Wnt Model M_E: Externally-Imposed Gradient, Static (M_{E1}) & Dynamic (M_{E2}) Updating

We investigated, under physiological conditions (i.e. wild-type cells) how crypt renewal dynamics are affected by contact inhibition and by the crosstalk between the Wnt, prescribed as an external gradient (models M_{E1} and M_{E2}), and the Hippo pathways by measuring the mitotic index (the percentage of cells undergoing mitosis at any given point in time, a proxy for the proliferative capacity of cells) and distribution of cell velocities over the crypt (Figure 5.14). The former is a measure of proliferative capacity, and the latter provides a snapshot of crypt motility. We varied

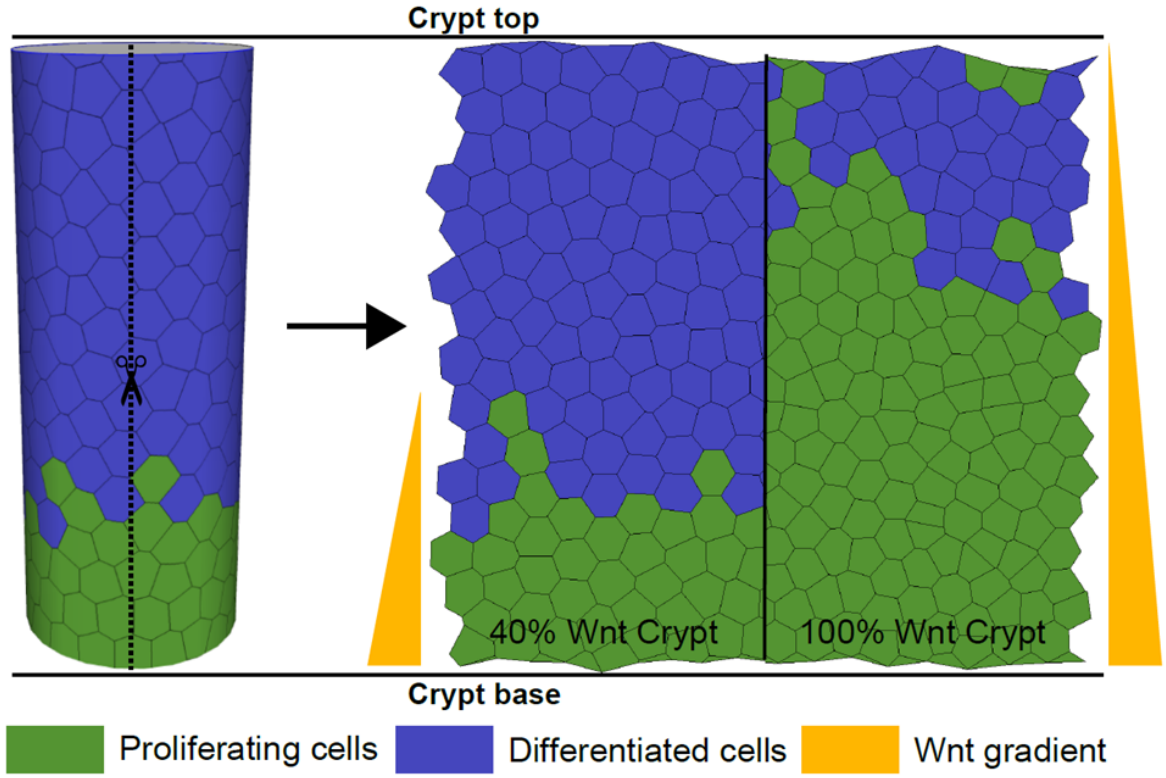


FIGURE 5.13. **Crypt ABM setup schematic.** Schematic diagram of the 3D to 2D projection method, 'unrolling' the cylindrical crypt to a planar domain with periodic boundary conditions. The right-hand panel also illustrates the imposed-gradient Wnt model M_{E1}, in the two considered cases of minimum Wnt level at 40% and 100% of the height of the crypt, and the resultant effect on the region in which cells proliferate.

two relevant parameters: the contact inhibition threshold, and the Wnt level within the crypt. The contact inhibition threshold – defined as the cell volume below which the YAP-P/β-catenin complex prevents progression through the cell cycle – is defined as a percentage of the equilibrium proliferating cell volume, and set to either 60% or 90%, (at 100% the majority of cells would be expected to be inhibited). The Wnt level – defined as the position in the crypt at which the external Wnt gradient reaches its minimum value – is set to either 40% or 100% of the height of the crypt (see Figure 5.13), as an attempt to account for the proliferative differences seen between crypts within the small [139, 140] and large intestine [159].

The effect on mitotic activity of varying Wnt level is that for a maximal Wnt level (Figure 5.14(a)) the proliferative 'niche' extends further up the crypt than for a lower Wnt level (Figure 5.14(b)); in the latter case there is a rapid drop-off of mitotic index at approximately 30% of the crypt height. This is expected when compared to the setup in Section 5.3 where we specified a

region encompassing the lower 35% of the crypt to be proliferative. Under all parameter settings, the main difference between the static (model M_{E1}) and dynamic (model M_{E2}) Wnt models (Figures 5.14a and b, solid and dashed lines respectively) is a more rapid drop off in the mitotic index using the latter approach, due to the fact that updating Wnt level dynamically causes cells to react instantly to changes in the external Wnt level as they move up the crypt, causing a more rapid cell-cycle cessation.

Varying the threshold volume at which cells are contact-inhibited (Figures 5.14a and b, red and blue lines), the volume-dependent prevention of β-catenin nuclear localisation, and hence the reduction in mitotic proportion, is more effective when cells reach the top of the proliferative niche where the Wnt signal decreases. These results match our single-cell simulations which suggested that the region over which contact inhibition is active is reduced with a smaller Wnt signal. Also, agent-based simulations indicate that inhibiting the cells at 90% of their equilibrium volume results in a reduction in mitotic activity of approximately 55% at the base of the crypt and of over 75% towards the top of the proliferative 'niche'. This suggests that CI within the crypt is as capable of affecting the proliferative activity of the crypt cells as changing the Wnt signalling update within the crypt.

A check of our new model is that the introduction of contact inhibition in a healthy wild-type crypt should not dramatically affect the qualitative behaviour of the proliferative dynamics when compared to previous models [49, 124, 173]. This is confirmed by the results shown in Figure 5.14, maintaining the proliferative 'niche' and expected mitotic activity. Simulations (Figures 5.14a and b) suggest that the differential effects of Wnt on proliferation dynamics can be greater than those of moderate CI effects; still, we can tune the CI threshold to cause a comparable reduction in crypt mitotic activity. The effect of increasing the threshold volume at which cells are contact inhibited, from 60% (blue) to 90% (red) of equilibrium volume, results in a reduction in mean crypt velocity (Figures 5.14c and d) of 63%, whilst decreasing the threshold for Wnt from 100% of the crypt to 40% of the crypt (Figures 5.14c and d, respectively) resulted in a reduction in mean crypt velocity of 53%. This qualitative similarity is due to the limited effect of contact inhibition in a wild-type crypt where cells are not, in general, significantly smaller than their equilibrium volume. Compared to the results shown in Section 5.3, in Figure 5.8 we see a similar reduction in the velocity of the cells within the crypt, there was an approximate reduction in velocity of 35% in the previous model compared to the 53% reduction in this case.

Overall, these results suggest that large-scale changes to cell proliferation are caused by altering Wnt signalling. However, Hippo signalling dependent contact inhibition is also reduces the proliferative activity within the crypt, with the largest effects seen in the 100% Wnt case, due to the larger number of proliferative cells, that can be affected by contact-inhibition. Considering the two Wnt hypotheses for the externally-imposed Wnt model – static (M_{E1}) and dynamic (M_{E2}) – the reduced mitotic activity present in the M_{E2} crypt is not in line with the mitotic activity seen in recent experimentation [169]. This suggests that the static (M_{E1}) crypt hypothesis may be

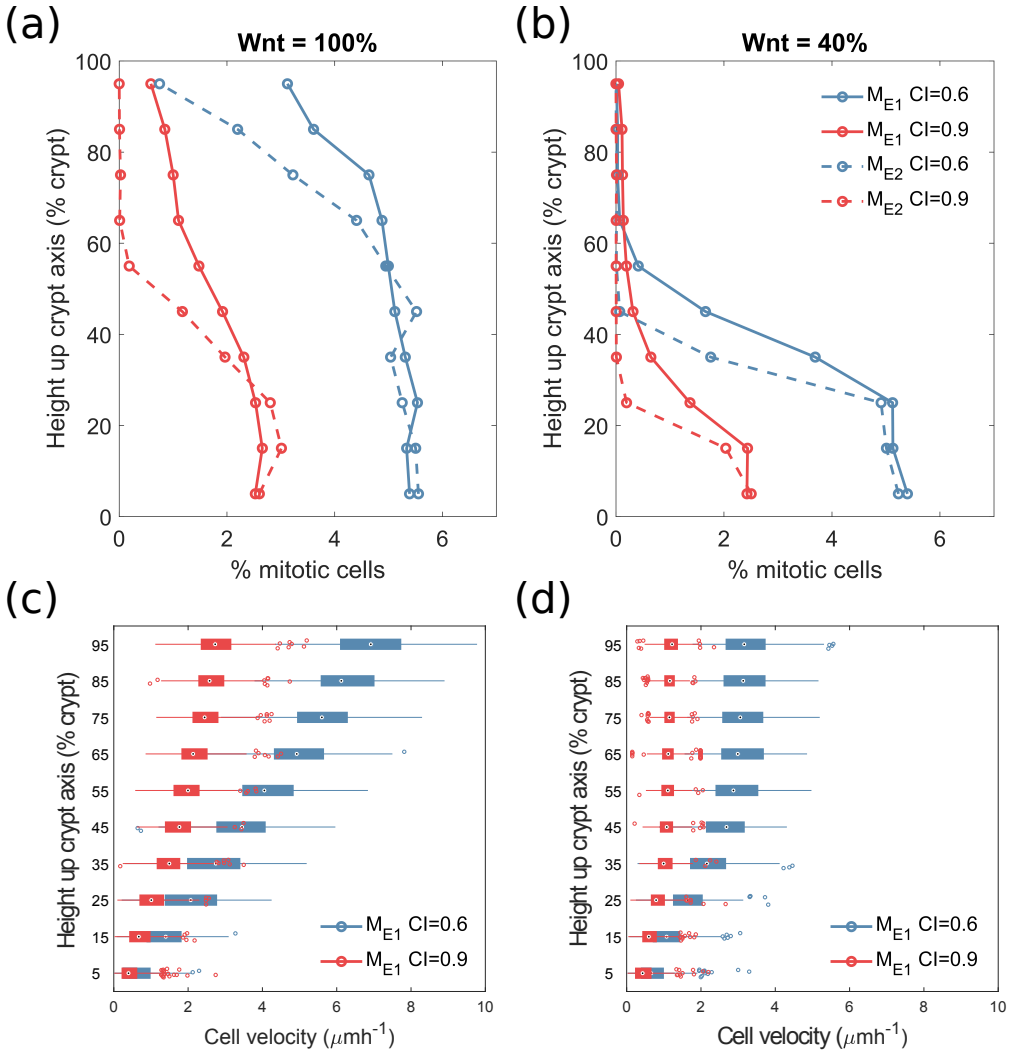


FIGURE 5.14. **Multiscale dynamics of Hippo and Wnt signalling, and Hippo-dependent contact-inhibition (CI), in wild-type crypt with imposed Wnt gradient (M_E). (a, b)** Effects of Hippo signalling-based CI on Wnt signalling-dependent cell-cycle measuring mitotic indices (proportion of cells undergoing mitosis at increasing intervals up the crypt). We considered both a static (M_{E1}, solid lines) and dynamic (M_{E2}, dashed lines) Wnt cell model (Wnt assigned at birth or continually updated, respectively), with Wnt signalling affecting 100% (a) and 40% (b) of the crypt. The different volumes at which cells undergo CI are indicated by the blue (CI at 60% of equilibrium volume) and red (CI at 90% of equilibrium volume) lines. **(c, d)** Velocity whisker plots for a crypt cell population using the static (M_{E1}) Wnt model, and Wnt signalling affecting 100% (c) or 40% (d) of the crypt, with CI occurring at 60% volume (blue) and 90% volume (red). The box plots indicate the median velocity of the cells at increasing heights up the crypt, with the box representing the 25th and 75th percentiles respectively. The whiskers extend to the most extreme velocities recorded over the course of the 150 repeated experiments.

more physiologically plausible; thus, in further simulations, only M_{E1} was considered. However, we would expect the effects of introduced CI to be the same in both hypotheses.

5.4.3.2 Wnt Model M_I: Cellular Division-Based Wnt

A prescribed fixed external gradient of Wnt is a feature of many computational models of the crypt dynamics, effectively prescribing a spatial proliferation threshold [49, 124, 173]. Given recent experimental results, that suggest an alternative division-based Wnt process [56], we modified our model to instead include division-dependent spreading of Wnt, carried independently by each cell, from a Wnt high reservoir at the base of the crypt. We investigated whether this approach could result in an emergent Wnt gradient in the crypt. We, therefore, set only cells at the base of the crypt to receive maximal Wnt signal, forming a "Wnt reservoir" whose size (as a proportion of the height of the crypt) can be varied as one of the parameters of the model. It has been shown that Wnt does not readily diffuse into surrounding cells [56], and we, therefore, neglected diffusion-based spreading of Wnt. Instead, Wnt is shared stochastically between daughter cells on division, so that daughter cells contain a proportion $0.5 \pm \epsilon$ of the mother cell's Wnt level at each division, where ϵ is a sample from a Normally distributed random variable with zero mean and standard deviation σ (set as a model parameter), appropriately truncated to ensure the proportion remains in the interval [0, 1].

Initial simulations were carried out with-noise independent Wnt allocation on division (i.e. σ=0), resulting in emergent Wnt gradients (Figures 5.15(a), (b)). Increasing the size of the Wnt reservoir from 10% to 20% of the crypt height reduces the steepness of the resultant Wnt gradient over the crypt domain, without affecting the mean Wnt density at the top of the crypt, so that it more closely resembles the imposed linear Wnt gradient of previous models.

We then introduced noise in the allocation of Wnt and repeated the wild-type crypt *in silico* experiments with 10% ($\sigma = 0.1$) and 20% ($\sigma = 0.2$) noise in the 10% and 20% Wnt reservoir crypts (Figures 5.15c, d). The mean Wnt gradient is not strongly affected by the introduction of this division noise. Mitotic index results (Figures 5.15(c) and (d)) suggest that for division-based spreading of Wnt with no noise (blue lines) there is a reduction followed by a sharp increase in mitotic activity in the approximate region where the first cell division occurs (25%/55% of the total height of the crypt, in case of a 10%/20% Wnt reservoir respectively). Introducing noise in the division of Wnt (red, green lines) this reduction and subsequent increase in mitotic activity is smoothed. This result suggests that the division of Wnt between daughter cells is not purely symmetric. Further, comparing results of externally imposed and division-based Wnt models (Figures 5.14(a), (b) and Figures 5.15 (c), (d), respectively), suggests that a crypt with a 10% Wnt reservoir is representative of a colonic crypt, while a 20% Wnt reservoir is more typical of a small intestinal crypt [139, 169]. The presence of Paneth cells in the small intestinal crypt could account for this required larger reservoir that we see in the experiments; conversely, an external Wnt signal that does not extend the full height of the crypt would also suffice in explaining a

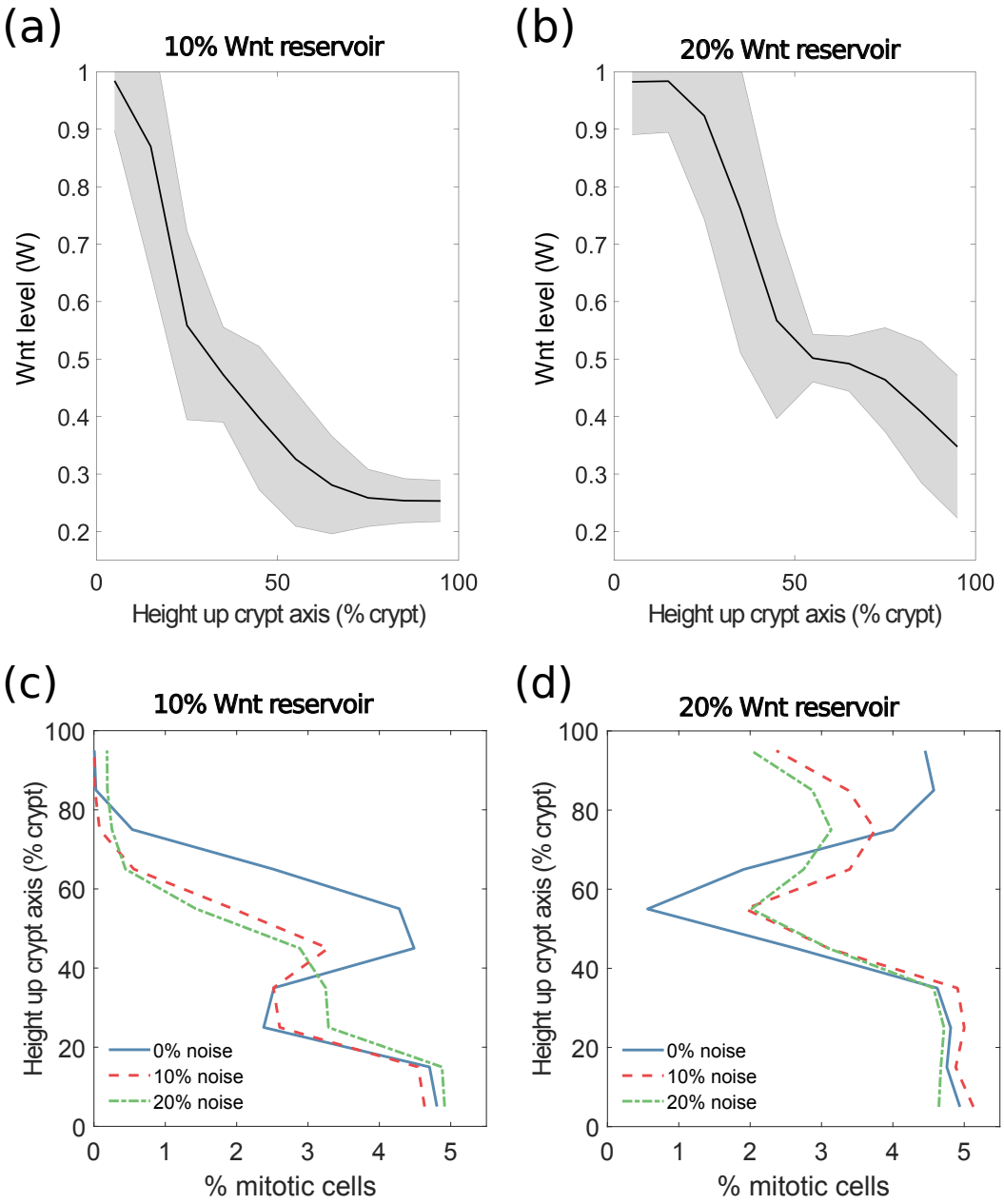


FIGURE 5.15. **Multiscale crypt dynamics using a cell division-based Wnt model (M_I). (a, b)** Wnt gradient through the crypt for the cell division-based Wnt model (M_I), with (a) 10%, and (b) 20% crypt-height Wnt reservoir. Solid line represents mean Wnt held, shaded region indicates the standard deviation amongst cells. **(c, d)** Mitotic activity, with asymmetric division of Wnt between daughter cells in crypts in (c) 10%, and (d) 20% crypt-height Wnt reservoir; blue, red and green lines represent symmetric division, 10% and 20% noise, respectively.

larger Wnt presence.

5.4.4 Introducing Mutations into the Crypt

Finally, we investigated the effects of CI on a dysplastic crypt; in this case, there can be a significant increase in cell number due to over-proliferation of mutant cells, resulting in a decrease in cell volume, at which point the effects of CI may play a more significant role. The mutant that we considered is a possible APC 'double mutant' [24], where there is total the disruption of the β-catenin destruction complex [93, 104] as well as disruption to Hippo signalling within the cell [92]; as a result, APC 'double mutant' cells hyper-proliferate throughout the crypt. The mutant has proliferative dynamics independent of both Wnt signal and cell volume. In what follows, we consider only the most realistic imposed external Wnt model (M_{E1}) and the alternative internal Wnt model (M_I). We conducted *in silico* experiments, varying the threshold volume at which Hippo signalling is active in the healthy cells, together with the size of the Wnt gradient (in M_{E1}) and of the Wnt reservoir (M_I); the results of insertion of a single affected APC 'double mutant' cell, in terms of mutant cell washout probabilities (i.e. the likelihood of healthy cells being able to remove the mutant from the crypt), are shown in Figure 5.16.

In the imposed Wnt gradient crypt (M_{E1}, Figure 5.16(a)), an increase of the Wnt signal from 40% to 100% of the crypt (Figure 5.16(a), blue and red lines respectively) resulted in only a small change to the washout probability at low levels of CI. Conversely, at the maximal CI volume threshold (90% of equilibrium volume), the washout probability increases from approximately 20% in the 40% Wnt case to 40% in the 100% Wnt cases, as the advantage of the mutant cells is reduced due to increase of proliferating healthy cells; recall that only wild-type cells are affected by CI, and that mutant cells proliferate independently of cell volume and Wnt level. This difference is statistically significant with a p-value <0.00001 at a significance level of 0.05. The increase in the level of CI (from 60% to 90% of equilibrium volume) reduces the washout probability by more than half in the 40% Wnt crypt (Figure 5.16(a), blue line). It is in this case where the advantage gained by a mutant cell from the combination of its increased proliferative capacity and CI is most significant, again with a statistical significance p-value <0.00001 at a significance level of 0.05. The reduction in washout probability is also comparable with previous models where additional advantage has been gained by the mutant cell by its increased adhesion to the crypt substrate and to the surrounding cells [124]. Of note, we did not model the altered adhesion of mutant cells, to specifically focus on signalling- and volume-dependent CI.

In the division-based Wnt model (M_I, Figure 5.16(b)), simulations of the same mutant takeover show that, for maximal CI volume (90% of equilibrium volume), an increase in the size of the Wnt reservoir from 10% to 20% of the height of the crypt (Figure 5.16(b), blue and red lines respectively) results in a large increase in the relative washout probability (from approximately 35% to 55%). Decreasing the volume threshold for CI resulted in a smaller decrease in the washout probability of that in the imposed-Wnt-gradient crypt with high Wnt (Figure 5.16,

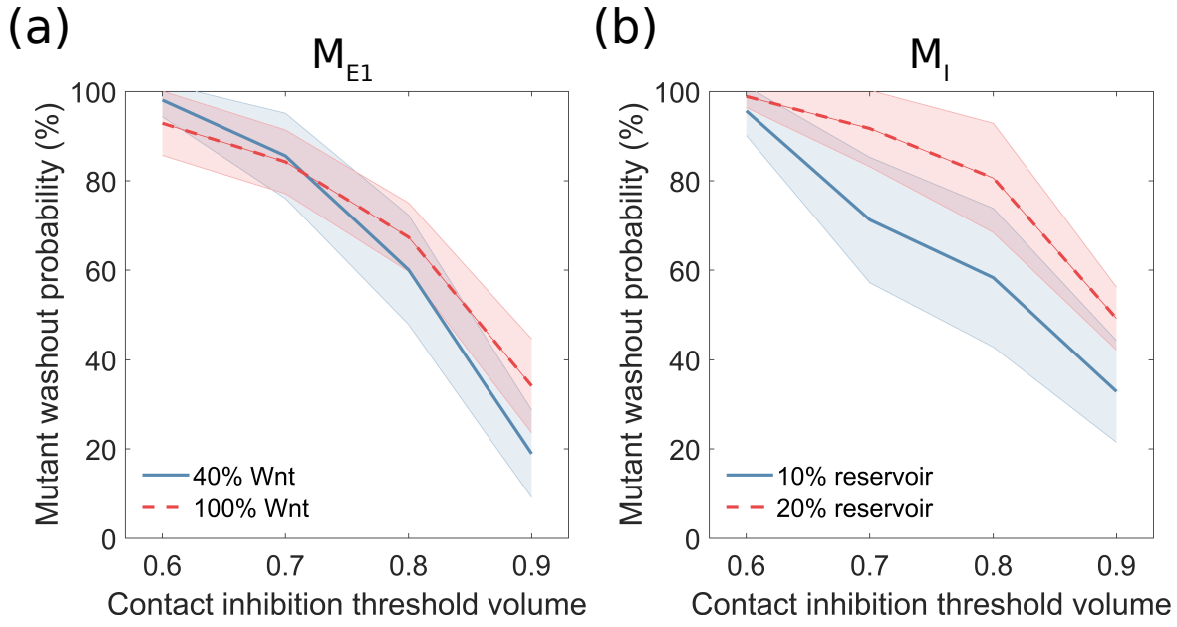


FIGURE 5.16. **Washout probabilities upon mis-regulated Wnt and Hippo signalling activity. (a,b)** Washout probabilities (the probability of the mutant cell being washed from the crypt) for APC double-mutant crypt, starting from a single mutant cell introduced at the base of the crypt (bottom 5% of crypt) using (a) externally-imposed gradient (M_{E1}) or (b) cell-division (M_I) Wnt models. Blue and red lines represent the low (40% Wnt threshold/10% Wnt reservoir) and high (100% Wnt threshold/20% Wnt reservoir) Wnt cases, respectively. Increasing the threshold for contact-inhibition reduces the probability of mutant washout. The shaded regions represent the 95% confidence interval for a binomial distribution with probability p, according to $p \pm z_{0.975}\sqrt{p(1-p)/n}$, where n is the number of simulations conducted, and $z_{0.975}$ is the 97.5 percentile of a standard normal distribution.

red lines), but similar behaviour for low-Wnt (Figure 5.16, blue lines). Overall, the simulated dysplastic dynamics qualitatively match those predicted by the imposed Wnt gradient model, with a reduction in washout probability resulting from a stronger level of CI, and an increase in washout probability in response to a higher level of Wnt in the crypt.

We can compare these washout probabilities with the takeover probabilities shown in Section 5.3, Figure 5.5. Doing so we see that at the lower level of CI (0.6) the probability of takeover in Figure 5.16 is approximately 5%, which is the same as the expected takeover probability in Figure 5.5 when $\alpha = 1$, which means there is no advantage to the mutant cell compared to the healthy cell. Increasing the CI threshold to the maximal value of 0.9 we see maximal takeover probabilities of approximately 80% (with a 40% Wnt crypt and M_{E1} Wnt model and 20% Wnt reservoir and M_I Wnt model), which compares with the approximately 60% takeover probability seen in Figure 5.5 $\alpha = 10$.

5.5 Discussion

Although the role of the Wnt signalling pathway in governing the proliferative dynamics of cells within intestinal crypts has been extensively studied, a comprehensive understanding of the mechanisms linking Wnt signalling to other relevant pathways, cell mechanical properties, and cell-cycle progression in health and disease is still lacking.

In this chapter, we considered how two ABMs of distinct complexity are able to represent the wild-type and mutant dynamics seen experimentally. In Section 5.3 we described an ABM of a colonic crypt which incorporated a Wnt threshold representation of the Wnt signalling pathway, linked to a fixed-length cell-cycle model, and included a volume threshold cell-cycle cessation model of CI. This work showed that incorporating CI into the model reduces the motility of cells, and cell-cell forcing, subsequently increasing the advantage between a healthy cell and an APC double mutant. Additionally, it appears that altering the threshold at which cells undergo CI alters the 'ribboning' behaviour seen when the mutant cell population count oscillates during the process of monoclonal conversion, effectively acting as a damper to the system. This reduces the time taken for mutant monoclonal conversion to occur and effectively increases the probability of this event.

In Section 5.4 we extended this to create a new multiscale computational model of crypt dynamics, incorporating a kinetic description of both Wnt/Hippo signalling and their effectors, as well as a gene network-based description of cell-cycle progression. This work showed that linking Wnt signalling at the point of cell division (static Wnt), as opposed to continually updating Wnt levels over the cells lifetime (dynamic Wnt), provides a more realistic representation of cell proliferation.

Previous models have assumed the existence of a fixed and externally-imposed gradient of Wnt within the crypt, with a maximal amount of Wnt found at the base, decreasing linearly with height. Based on recent experimental observations, we developed a model that assumes Wnt is internally held by each cell, supplied initially only to a portion of them in a reservoir at the base of the crypt. We showed that subsequent division events create an emergent gradient of Wnt similar to that imposed by existing Wnt models. Noise in the Wnt allocation to daughter cells following mitosis did not significantly affect the overall Wnt gradient but smoothed the distribution of mitotic activity within the crypt. Increasing the size of the Wnt reservoir controlled the properties of the Wnt gradient, in terms of its steepness and linearity. Furthermore, our results suggest that the spatial localisation of the Wnt reservoir can affect the crypt's dynamics, both wild-type and dysplastic, as much as the total amount of Wnt available.

In order to elucidate the mechanisms underpinning cellular CI within the crypt, we extended a Wnt regulatory network model to additionally include the role of Hippo signalling in phosphorylating YAP/TAZ; the latter has been suggested to trap β-catenin within the cytoplasm and at the extracellular matrix, preventing nuclear accumulation and subsequent expression of Wnt target genes linked to cell-cycle progression [125]. Our extended regulatory network is sensitive

to a reduction in nuclear accumulation of β-catenin; single-cell model simulations showed that Hippo signalling can reduce nuclear β-catenin accumulation, and cause the rapid cessation of the cell cycle, matching experimental observations [92]. Agent-based simulations including the Wnt/Hippo signalling coupling, with the levels of Hippo signalling further linked to cell volume, validated the single-cell analysis: the length of the cell cycle increased, and eventually ceases, as the cell volume decreases. We found that cells with low Wnt levels are more susceptible to CI: as the level of unphosphorylated β-catenin within the cell is reduced, the amount needed to bind to YAP-P to prevent nuclear accumulation and subsequent transcription of target genes linked to proliferation is reduced.

To investigate the effect of mutations within the crypt, we performed agent-based simulations (using both an external-gradient Wnt model with static Wnt, and the internal division-based Wnt model) with an APC 'double mutant' cell introduced into the crypt. The simulated mutation caused altered destruction complex kinetics in the Wnt signalling model, as well as disruption of the Hippo signalling module. Increasing the amount of Wnt within the crypt did not affect the washout probability of the mutant cells at low levels of CI, with the advantage of the mutant cells not significantly reduced by the increase in proliferation caused by Wnt upregulation. The main difference arose from increasing the volume threshold for CI in the cells, with the largest decrease in the washout probability seen at a volume threshold of 90% of the healthy cell equilibrium. The effect on washout probability and cell velocity of this mutant was similar to that of increased adhesion observed previously [60], shown in Figure 5.5, with maximal probabilities of monoclonal conversion of 50% in the previous model and of 55% in our extended model of the crypt. Critically, moving from the external-gradient to the internally held and division-based Wnt model did not qualitatively affect the increased mutant advantage, further suggesting cell division as a plausible mechanism for the experimentally-observed Wnt gradient [56, 153].

CHAPTER 6

DISCUSSION

This thesis has aimed to show the ability of mathematical and computational models to replicate stem cell dynamics *in vitro* and *in vivo*, to effectively capture different culture techniques and methods, and to inform future experimental research, as well as to direct us towards avenues for future modelling development. Whilst each of the models discussed within this thesis has involved a degree of simplification, they have all accurately captured the dynamics of the individual stem cell types, and provided an insight into the mechanisms governing their most important behaviours, linked to pluripotency, proliferation and differentiation. In this chapter, in Section 6.1 we summarise and discuss the key results and knowledge gained from our investigations into human haematopoietic stem cells (Chapter 3), mouse embryonic stem cells (Chapter 4) and intestinal stem cells (Chapter 5), whilst providing a roadmap for future research into these systems as well as the general effectiveness of mathematical and computational modelling of stem cell dynamics. In Section 6.2 we suggest possible future avenues of research.

6.1 Key Findings

6.1.1 Human Haematopoietic Stem Cells (Chapter 3)

In Chapter 3 we investigated human haematopoietic stem cell (HHSC) dynamics in culture, modelling the differentiation and proliferation processes during a key transitional step along the haematopoietic stem cell lineage tree. We were interested in quantifying how the application of a fundamental hormone involved in stimulating red blood cell production *in vivo*, erythropoietin (epo), affects the differentiative and proliferative dynamics of two key cell types, common myeloid progenitor (CMP) and megakaryocyte-erythroid progenitor (MEP), in cell cultures *in vitro*. Whilst it was known that epo was essential for the production of mature erythroid cells *in vivo* [102],

promoting both proliferation and survival, the requirements for *in vitro* cell culture were unclear. As such, experimentalists typically applied epo from the start of the cell culturing process, something which is potentially extremely costly and potentially unnecessary.

To achieve this, we developed a mathematical model that captured the differential effects of epo on proliferation and differentiation of HHSCs. By extending the Smith-Martin model [155], a formalism which uses delay differential equations (DDEs) to describe cell-cycle progression and thus cell proliferation, we were able to simultaneously describe the differentiation of HHSCs by modelling both CMP and MEP populations and incorporating the differentiation step in the model.

Using such an extended DDE model for cell proliferation and differentiation, we were able to quantitatively fit the growth dynamics of HHSC cultures, with a root mean square error (RMSE) between the experimental data and our model of approximately 5%, in line with the variation found between experimental data sets.

Fitted parameters showed that the proliferation rate of CMP cells increases with each generation, whilst the proliferation rate of MEP cells remained relatively constant, with a standard deviation of less than 5% of the mean rate. The increased proliferation rate of CMP, combined with the constant proliferation of MEP, was unexpected, as it suggests that the forming of a maximal MEP colony has a greater dependency on the proliferation of the less mature cell type and not their own cell type.

As a result of our model fitting we were also able to show that, unexpectedly, the maturation of cells from CMP to MEP cells is weighted towards the early stage of culture, when the cells are younger. We found that the majority of maturation occurs before the midpoint of the culture experiment (i.e. day 2); as such, our fitted maturation rate for CMP cells is maximal at the start of the experiment and follows a logistic curve dropping to a minimal value by the end of the time-course. The result of this maturation is that by the end of the experiments (i.e. day 4) approximately 75% of cells have matured from generation 0 of CMP to MEP.

Following on from fitting our model to wild-type HHSC culture experimental data, we used our framework to analyse the effects of epo on the cultured population and to determine the culture requirements of HHSCs. We showed that, from the initial stages of megakaryocyte/erythroid lineage commitment, cells are epo responsive and proliferation is slower in the absence of epo. By disentangling the independent contributions of proliferation and maturation to population dynamics, the model allowed a high-resolution quantification of proliferation rate that revealed the sensitivity of MEP to epo. It was clear that there is a distinct increase in the number of MEP cells present when cultured with epo; we were able to quantify the effects of epo on proliferation and differentiation of CMP and MEP cells. We found, on average, a 11.04($\pm$4.88)% increase in the proliferation of CMP cells, and a 27.93($\pm$11.49)% increase in the proliferation of MEP.

In validating our model, and its capability to replicate the culture dynamics of HHCSs, we used the parameters fitted from the 4-day culture experiments to extrapolate the cell culture

dynamics for an additional 8 days. We were able to show that the predicted proliferation rates in the presence of epo were able to faithfully reproduce the experimental data. This confirms the power of the model to describe erythroid cultures and to predict changes in the population dynamics. However, when simulating the population in the absence of epo, we observed a marked change occurring after 5 days of culture. There appeared to be a more pronounced reduction in the size of the experimental population than we would predict with our model. This suggests that there are two stages of MEP cultures, each with different sensitivity to epo, an early stage (Days 1–4) with a small increase in proliferation rate in response to epo, and a later stage (Days 4–8), which is increasingly dependent on epo for proliferation.

6.1.2 Mouse Embryonic Stem Cells (Chapter 4)

In Chapter 4 we used agent-based modelling to replicate culturing experiments of mouse embryonic stem cells (mESCs), and to couple single cell dynamics to cell proliferation. Recent work on the gene-regulatory network governing pluripotency of single mESCs [68] was able to account for mESCs cultured in two different culture media, serum/LIF and 2i/LIF, replicating the bistability of a key pluripotency gene, Nanog. We designed an agent-based model which incorporated this gene-regulatory network and used it to determine whether the results of the single-cell analysis hold at a population level, by linking the underlying GRN to the cell cycle. We investigated the role of Nanog in regulating cell-cycle behaviours, as suggested by previous studies [83], and additionally extended the model to include MycN, a gene whose role in regulating cell proliferation and differentiation is less thoroughly studied but potentially significant [150].

By using Nanog as a cell-cycle regulator, we were able to quantitatively capture the distribution of Nanog seen experimentally; however, when we considered the population growth dynamics, there was a significant (80% greater than expected) overestimation of the population count of the cells when we switched from the serum/LIF to the 2i/LIF medium.

In order to capture both the gene concentration distributions as well as the growth dynamics, we considered the role of the gene MycN as a regulator of cell-cycle behaviour. By linking the concentration of MycN present in individual mESCs to their proliferative capacity within the population, we were able to capture the fold change characteristics of the cell culture in both serum/LIF and in 2i/LIF, as well as the distribution of Nanog within the population. Our fitted agent-based model was able to produce simulated populations that matched qualitatively (by capturing the expected reduced proliferation in the switch from serum/LIF to 2i/LIF) the experimental data.

Additionally, we were able to reproduce the sorting experiment results that were suggested in the equation-based modelling work done previously, with the *in-silico* sorting experiment results recapitulating the expected distributions. Considering the effects of MycN on the proliferative behaviour of the cells has allowed us to better describe the link between pluripotency and cell proliferation, which might be relevant for the design of superior protocols for embryonic stem cell

cultures.

6.1.3 Intestinal Crypt Stem Cells (Chapter 5)

In Chapter 5 we focused on the role of signalling pathways in the maintenance of healthy intestinal crypts. Although the role of the Wnt signalling pathway in governing the proliferative dynamics of cells within intestinal crypts has been extensively studied, a comprehensive understanding of the mechanisms linking Wnt signalling to other relevant pathways, cell mechanical properties, and cell-cycle progression in health and disease was still lacking. Additionally, we were interested in further understanding the signalling pathways governing contact inhibition (CI), where the cessation of the cell cycle can be brought on by mechanistic cues within the crypt system.

We considered how two agent-based models (ABMs) are able to represent the wild-type and mutant dynamics seen experimentally. We first described an ABM of a colonic crypt which incorporated a threshold-based representation of the Wnt signalling pathway, linked to a fixed-length cell-cycle model, and included a volume threshold cell-cycle cessation model of CI. We showed that incorporating CI into a crypt model is able to reduce the motility of cells by decreasing cell-cell forcing and subsequently increasing the advantage between a healthy cell and an APC double mutant. Additionally, it appears that varying the threshold at which cells undergo CI alters the 'ribboning' behaviour seen when the mutant cell population count oscillates during the process of monoclonal conversion, effectively acting as a damper to the system. This reduces the time taken and increased the likelihood of, mutant monoclonal conversion to occur.

We extended on this simple representation of CI to create a new multiscale computational model of crypt dynamics, incorporating a network description of both Wnt/Hippo signalling and their effectors, as well as a network description of cell-cycle progression. We showed that linking Wnt signalling at the point of cell division (static Wnt), as opposed to continually updating Wnt levels over the cells' lifetimes (dynamic Wnt), provides a more realistic representation of cell proliferation.

Previous models assumed the existence of a fixed and externally-imposed gradient of Wnt within the crypt. Based on recent experimental observations, we developed a model that assumes Wnt is internally held by each cell, supplied initially only to cells in a reservoir at the base of the crypt. We showed that subsequent division events create an emergent gradient of Wnt similar to that imposed by existing Wnt models. We found that introducing noise in the Wnt allocation to daughter cells following mitosis does not significantly affect the overall Wnt gradient, but it does manage to smooth the distribution of mitotic activity within the crypt. Additionally, increasing the size of the Wnt reservoir altered the properties of the Wnt gradient, in terms of its steepness and linearity. Furthermore, our results suggest that the spatial localisation of the Wnt reservoir can affect the crypt's dynamics, both wild-type and dysplastic, as much as the total amount of Wnt available.

In order to further understand mechanisms underpinning cellular CI within the crypt, we extended a Wnt regulatory network model to additionally include the role of Hippo signalling in phosphorylating YAP/TAZ; the latter has been suggested to trap β-catenin within the cytoplasm and at the extracellular matrix, preventing nuclear accumulation and subsequent expression of Wnt target genes linked to cell-cycle progression [125]. We found that our extended regulatory network is sensitive to a reduction in nuclear accumulation of β-catenin: our single-cell model simulations showed that Hippo signalling can reduce nuclear β-catenin accumulation, and cause the rapid cessation of the cell-cycle, matching experimental observations [92]. Agent-based simulations including the Wnt/Hippo signalling coupling, with the levels of Hippo signalling further linked to cell volume, validated the single-cell analysis: the length of the cell-cycle increases, and eventually ceases, as the cell volume decreases. Cells with low Wnt levels are more susceptible to CI; as the level of unphosphorylated β-catenin within the cell is reduced, the amount needed to bind to YAP-P to prevent nuclear accumulation and subsequent transcription of target genes linked to proliferation is reduced.

To investigate the effect of mutations within the crypt, we performed agent-based simulations (using both the external-gradient Wnt model with static Wnt and the internal division-based Wnt model) with an APC 'double mutant' cell introduced into the crypt. The simulated mutation caused altered destruction complex kinetics in the Wnt signalling pathway, as well as disruption of the Hippo signalling module. Increasing the amount of Wnt within the crypt did not affect the washout probability of the mutant cells at low levels of CI, with the advantage of the mutant cells not significantly reduced by the increase in proliferation caused by Wnt up-regulation. The main difference arose from increasing the volume threshold for CI in the cells, with the largest decrease in the washout probability seen at a volume threshold of 90% of the healthy cell equilibrium. The effect on washout probability and cell velocity of this mutant was similar to that of increased adhesion observed previously [60]. Critically, moving from the external-gradient to the internally held and division-based Wnt model did not qualitatively affect the increased mutant advantage, further suggesting cell division as a plausible mechanism for the experimentally-observed Wnt gradient [56, 153].

6.2 Future Research

The results from our investigation into HHSCs provide some key insights with specific practical applications. We were able to show that the incorporation of epo during the early stages of cell culture is necessary, with an 11% increase in the proliferation of CMP and a 28% increase in the proliferation of MEP cells when cultured in the presence of epo as compared to those cultured without epo. However, our work also uncovered a potential dual stage response to epo, with a greater reduction in proliferation seen when cells are cultured without epo, beyond the first 4 days of culture. Our extrapolated model extension predicted a population over 6.5 times greater

than at the end of the 8-day of culture without epo. This dual stage response was unexpected and additionally not clear from the experimental data whole population counts. The implications of this dual stage response are that culturing HHSCs may benefit from a staggered instead of constant application of epo, with an increased application at a later stage of the culture. In considering this staggered application of epo, the use of mathematical models such as that developed in this thesis would be useful in determining the optimal application of important hormones during the course of the cell culturing. These methods could be applied systematically to data collected from varying application times and quantities of epo to determine the optimal application; alternatively, if we refitted the model to the extrapolated dataset, we could potentially capture the changing proliferative dynamics of the cell between the two response stages. The use of mathematical modelling to determine the effects of varying hormone concentration, or administration times, could lead to potential reductions in cost whilst simultaneously leading to an increase in the number of healthy cells that can be cultured within a shorter period of time.

Considering our work on intestinal crypts, the results suggest the possibility of combined disease treatment, targeting both the Wnt and Hippo pathways. There are also several possible avenues for further study, to better understand the dynamics of the crypt. In modelling the sub-cellular dynamics of crypt epithelial cells and crypt organisation, we did not consider the role of additional pathways, e.g. the Notch signalling pathway, which has been shown to interact with the Wnt signalling pathway to govern cell fate specification and cell differentiation via the Notch target Hes1, and the protein Hath 1 [99]. This Notch target is also regulated by β-catenin, a protein involved in the Wnt signalling pathway, and as such there is likely crosstalk occurring between the two pathways regulating cell proliferation as well as cell fate specification. In the division-based Wnt model, we assumed a reservoir at the base of the crypt within which cells can uptake Wnt, as a simplification of the presence of Wnt secreting cells at the crypt base. A more detailed model should explicitly consider such cells, and Wnt secretion, in addition to possible uptake from an external reservoir. In particular, for small intestinal crypts, the aim would be to extend the model to incorporate Paneth cells, secretory differentiated cells which reside at the base of the crypt and secrete stem cell maintenance factors, including Wnt [57]. An extension to our model could combine the Notch and Wnt signalling pathway models to allow dynamic differentiation and regulation of this Paneth cell based Wnt reservoir.

Our model also simplifies the Hippo-based sequestration of β-catenin within the cell; one valuable extension would be to explore the extent to which Hippo signalling either traps β-catenin within the cytoplasm or actively transports it to the extracellular matrix, in turn causing changes to the adhesive behaviour of the cells. By incorporating the effects of Hippo signalling affecting cell-cell adhesion we could compare more readily with previous modelling attempts investigating the role of adhesion in crypt maintenance and monoclonal conversion [124].

One potentially exciting extension of the work would be to investigate the mechanism for crypt formation and maintenance in intestinal organoids. Intestinal organoids represent a significant

opportunity to investigate crypts *in vitro*; while computational models exist in the literature [105], they do not consider the specific differentiative path and formation of the Wnt gradient that is found within crypts *in vivo*. Previous organoid modelling attempts have focused heavily on the differential effects of cell stiffness on the formation of crypts through the buckling of an epithelial surface [46]. By introducing Wnt-secreting cells, and combining them with signalling network dependent cell-cycle models such as those considered in Chapter 5, it could be possible to create a self-sustaining organoid growth model which mimics the *in vivo* dynamics of intestinal crypts. Our models already can form a Wnt gradient in a single crypt; by also introducing Wnt secretory cells it would be possible to grow organoids with proliferative dynamics seen in intestinal crypts, with a Paneth cell niche at the base of the crypt. The combination of Paneth cell-governed proliferative dynamics, with the differential stiffness Paneth cells at the base compared to those of the semi-differentiated cells forming the crypt walls, could lead to a significantly improved understanding of the dynamics during the formation of intestinal crypts.

Whilst this thesis has focused on three particular stem cell population types, these methods have the potential to be applied to other cell populations. Attempting to discern intricate interactions between individual cell dynamics and subcellular interactions from whole population level data is not trivial, and so the use of mathematical and computational modelling techniques such as those described in this thesis allows us to further understand complex interplay and multiscale behaviours in health and disease.

BIBLIOGRAPHY

[1] M. ABERCROMBIE, *Contact inhibition in tissue culture*, In Vitro, 6 (1970), pp. 128–142.

[2] E. ABRANCHES, E. BEKMAN, AND D. HENRIQUE, *Generation and characterization of a novel mouse embryonic stem cell line with a dynamic reporter of Nanog expression*, PLoS One, 8 (2013), p. e59928.

[3] E. ABRANCHES, A. M. V. GUEDES, M. MORAVEC, H. MAAMAR, P. SVOBODA, A. RAJ, AND D. HENRIQUE, *Stochastic NANOG fluctuations allow mouse embryonic stem cells to explore pluripotency*, Development, 141 (2014), pp. 2770–2779.

[4] J. D. AITCHISON AND T. GALITSKI, *Inventories to insights*, 2003.

[5] A. ATALA, *Re: Collective and single cell behavior in epithelial contact inhibition*, Journal of Urology, 188 (2012), pp. 1396–1397.

[6] A. M. BAKER, B. CERESER, S. MELTON, A. G. FLETCHER, M. RODRIGUEZ-JUSTO, P. J. TADROUS, A. HUMPHRIES, G. ELIA, S. A. MCDONALD, N. A. WRIGHT, B. D. SIMONS, M. JANSEN, AND T. A. GRAHAM, *Quantification of crypt and stem cell evolution in the normal and neoplastic human colon*, Cell Reports, 8 (2014), pp. 940–947.

[7] H. BANKS AND W. CLAYTON THOMPSON, *Mathematical Models of Dividing Cell Populations: Application to CFSE Data*, Mathematical Modelling of Natural Phenomena, 7 (2012), pp. 24–52.

[8] H. BANKS AND W. THOMPSON, *A division-dependent compartmental model with cyton and intracellular label dynamics*, International Journal of Pure and Applied Mathematics, 77 (2012), pp. 119–147.

[9] M. BARBERIS, E. KLIPP, M. VANONI, AND L. ALBERGHINA, *Cell size at S phase initiation: An emergent property of the G1/S network*, PLoS Computational Biology, 3 (2007), pp. 649–666.

[10] N. BARKER, R. A. RIDGWAY, J. H. VAN ES, M. VAN DE WETERING, H. BEGTHEL, M. VAN DEN BORN, E. DANENBERG, A. R. CLARKE, O. J. SANSOM, AND H. CLEVERS, *Crypt stem cells as the cells-of-origin of intestinal cancer*, Nature, 457 (2009), pp. 608–611.

[11] E. R. BARRY AND F. D. CAMARGO, *The Hippo superhighway: Signaling crossroads converging on the Hippo/Yap pathway in stem cells and development*, Current Opinion in Cell Biology, 25 (2013), pp. 247–253.

[12] M. BASAN, T. IDEMA, M. LENZ, J. F. JOANNY, AND T. RISLER, *A Reaction-Diffusion Model of the Cadherin-Catenin System: A Possible Mechanism for Contact Inhibition and Implication for Tumorigenesis.*, Biophysical Journal, 98 (2010), pp. 2770–2779.

[13] J. G. BETTS, P. DESAIX, E. JOHNSON, J. E. JOHNSON, O. KOROL, D. KRUSE, B. POE, J. A. WISE, M. WOMBLE, AND K. A. YOUNG, *Anatomy and Physiology*, 2016.

[14] M. BOIANI AND H. R. SCHÖLER, *Regulatory networks in embryo-derived pluripotent stem cells*, 2005.

[15] V. BOLÓS, J. GREGO-BESSA, AND J. L. DE LA POMPA, *Notch signaling in development and cancer*, 2007.

[16] B. M. BOMAN, J. Z. FIELDS, O. BONHAM-CARTER, AND O. A. RUNQUIST, *Computer Modeling Implicates Stem Cell Overproduction in Colon Cancer Initiation Advances in Brief Computer Modeling Implicates Stem Cell Overproduction in Colon*, Cancer, 61 (2001), pp. 8408–8411.

[17] C. BOSETTI, F. LEVI, V. ROSATO, P. BERTUCCIO, F. LUCCHINI, E. NEGRI, AND C. LA VECCHIA, *Recent trends in colorectal cancer mortality in Europe*, International Journal of Cancer, 129 (2011), pp. 180–191.

[18] F. A. BROOK AND R. L. GARDNER, *The origin and efficient derivation of embryonic stem cells in the mouse*, Proceedings of the National Academy of Sciences, 94 (1997), pp. 5709–5712.

[19] G. BÜSCHE AND H. KREIPE, *Blood and Bone Marrow Pathology*, Churchill Livingstone, 2011.

[20] S. BUSENBERG AND B. TANG, *Mathematical models of the early embryonic cell cycle: the role of MPF activation and cyclin degradation*, Journal of Mathematical Biology, 32 (1994), pp. 573–596.

[21] P. BUSKE, J. GALLE, N. BARKER, G. AUST, H. CLEVERS, AND M. LOEFFLER, *A comprehensive model of the spatio-temporal stem cell and tissue organisation in the intestinal crypt*, PLoS Computational Biology, 7 (2011).

[22] ——, *A comprehensive model of the spatio-temporal stem cell and tissue organisation in the intestinal crypt*, PLoS Computational Biology, 7 (2011).

[23] P. CAHAN AND G. Q. DALEY, *Origins and implications of pluripotent stem cell variability and heterogeneity*, Nature Reviews Molecular Cell Biology, 14 (2013), pp. 357–368.

[24] J. CAI, A. MAITRA, R. A. ANDERS, M. M. TAKETO, AND D. PAN, *β -Catenin destruction complex- independent regulation of Hippo ‚Äì YAP signaling by APC in intestinal tumorigenesis*, Genes & development, (2015), pp. 1–14.

[25] J. CAI, N. ZHANG, Y. ZHENG, R. F. DE WILDE, A. MAITRA, AND D. PAN, *The Hippo signaling pathway restricts the oncogenic potential of an intestinal regeneration program*, Genes and Development, 24 (2010), pp. 2383–2388.

[26] F. D. CAMARGO, S. GOKHALE, J. B. JOHNNIDIS, D. FU, G. W. BELL, R. JAENISCH, AND T. R. BRUMMELKAMP, *YAP1 Increases Organ Size and Expands Undifferentiated Progenitor Cells*, Current Biology, 17 (2007), pp. 2054–2060.

[27] H. CHENG AND C. P. LEBLOND, *Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine V. Unitarian theory of the origin of the four epithelial cell types*, American Journal of Anatomy, (1974).

[28] S. H. CHESHIER, S. J. MORRISON, X. LIAO, AND I. L. WEISSMAN, *In vivo proliferation and cell cycle kinetics of long-term self-renewing hematopoietic stem cells*, Proceedings of the National Academy of Sciences, 96 (1999), pp. 3120–3125.

[29] H. J. CHO, H. JÖNSSON, K. CAMPBELL, P. MELKE, J. W. WILLIAMS, B. JEDYNAK, A. M. STEVENS, A. GROISMAN, AND A. LEVCHENKO, *Self-organization in high-density bacterial colonies: Efficient crowd control*, PLoS Biology, 5 (2007), pp. 2614–2623.

[30] C. M. CLANCY, *The Need for*. \url{https://www.blood.co.uk/why-give-blood/the-need-for-blood/}, 2006.

[31] H. CLEVERS AND R. NUSSE, *Wnt / β-catenin signaling and disease*, 2012.

[32] N. COLLIER AND M. NORTH, *Parallel agent-based simulation with Repast for High Performance Computing*, Simulation, 89 (2013), pp. 1215–1235.

[33] J. COOPER, R. J. SPITERI, AND G. R. MIRAMS, *Cellular cardiac electrophysiology modeling with Chaste and CellML*, Frontiers in Physiology, 6 (2015).

[34] C. CROSNIER, D. STAMATAKI, AND J. LEWIS, *Organizing cell renewal in the intestine: Stem cells, signals and combinatorial control*, 2006.

[35] J. H. CUMMINGS, E. W. POMARE, H. W. J. BRANCH, C. P. E. NAYLOR, AND G. T. MACFARLANE, *Short chain fatty acids in human large intestine, portal, hepatic and venous blood*, Gut, 28 (1987), pp. 1221–1227.

[36] S. DALTON, *Linking the Cell Cycle to Cell Fate Decisions*, 2015.

[37] Y. DAVIT, J. M. OSBORNE, H. M. BYRNE, D. GAVAGHAN, AND J. PITT-FRANCIS, *Validity of the Cauchy-Born rule applied to discrete cellular-scale models of biological tissues*, Physical Review E - Statistical, Nonlinear, and Soft Matter Physics, 87 (2013).

[38] R. J. DE BOER, V. V. GANUSOV, D. MILUTINOVIĆ, P. D. HODGKIN, AND A. S. PERELSON, *Estimating lymphocyte division and death rates from CFSE data*, Bulletin of Mathematical Biology, 68 (2006), pp. 1011–1031.

[39] G. DE MATTEIS, A. GRAUDENZI, AND M. ANTONIOTTI, *A review of spatial computational models for multi-cellular systems, with regard to intestinal crypts and colorectal cancer development*, Journal of Mathematical Biology, 66 (2013), pp. 1409–1462.

[40] E. K. DEENICK, A. V. GETT, AND P. D. HODGKIN, *Stochastic Model of T Cell Proliferation: A Calculus Revealing IL-2 Regulation of Precursor Frequencies, Cell Cycle Time, and Survival*, The Journal of Immunology, 170 (2003), pp. 4963–4972.

[41] G. DENNIS, B. T. SHERMAN, D. A. HOSACK, J. YANG, W. GAO, H. LANE, AND R. A. LEMPICKI, *DAVID: Database for Annotation, Visualization, and Integrated Discovery*, Genome Biology, 4 (2003), p. R60.

[42] J. M. DESESSO AND C. F. JACOBSON, *Anatomical and physiological parameters affecting gastrointestinal absorption in humans and rats*, Food and Chemical Toxicology, 39 (2001), pp. 209–228.

[43] A. DI GARBO, M. D. JOHNSTON, S. J. CHAPMAN, AND P. K. MAINI, *Variable renewal rate and growth properties of cell populations in colon crypts*, Physical Review E - Statistical, Nonlinear, and Soft Matter Physics, 81 (2010).

[44] J. DONG, G. FELDMANN, J. HUANG, S. WU, N. ZHANG, S. A. COMERFORD, M. GAYYED, R. A. ANDERS, A. MAITRA, AND D. PAN, *Elucidation of a Universal Size-Control Mechanism in Drosophila and Mammals*, Cell, 130 (2007), pp. 1120–1133.

[45] D. DRASDO AND S. HÖHME, *A single-cell-based model of tumor growth in vitro : monolayers and spheroids*, Physical Biology, 2 (2005), pp. 133–147.

[46] S. J. DUNN, P. L. APPLETON, S. A. NELSON, I. S. NÄTHKE, D. J. GAVAGHAN, AND J. M. OSBORNE, *A two-dimensional model of the colonic crypt accounting for the role of the*

basement membrane and pericryptal fibroblast sheath, PLoS Computational Biology, 8 (2012).

[47] S. J. DUNN, A. G. FLETCHER, S. J. CHAPMAN, D. J. GAVAGHAN, AND J. M. OSBORNE, *Modelling the role of the basement membrane beneath a growing epithelial monolayer*, Journal of Theoretical Biology, 298 (2012), pp. 82–91.

[48] S. J. DUNN, I. S. NÄTHKE, AND J. M. OSBORNE, *Computational models reveal a passive mechanism for cell migration in the crypt*, PLoS ONE, 8 (2013), pp. 1–18.

[49] S.-J. DUNN, J. M. OSBORNE, P. L. APPLETON, AND I. NÄTHKE, *Combined changes in Wnt signaling response and contact inhibition induce altered proliferation in radiation-treated intestinal crypts*, Molecular Biology of the Cell, 27 (2016), pp. 1863–1874.

[50] S. DUPONT, L. MORSUT, M. ARAGONA, E. ENZO, S. GIULITTI, M. CORDENONSI, F. ZANCONATO, J. LE DIGABEL, M. FORCATO, S. BICCIATO, N. ELVASSORE, AND S. PICCOLO, *Role of YAP/TAZ in mechanotransduction*, Nature, 474 (2011), pp. 179–184.

[51] E. DZIERZAK AND S. PHILIPSEN, *Erythropoiesis: Development and differentiation*, Cold Spring Harbor Perspectives in Medicine, 3 (2013).

[52] P. EDELMAN, G. VINCI, J. L. VILLEVAL, W. VAINCHENKER, A. HENRI, R. MIGLIERINA, P. ROUGER, J. REVIRON, J. BRETON-GORIUS, AND C. SUREAU, *A monoclonal antibody against an erythrocyte ontogenic antigen identifies fetal and adult erythroid progenitors.*, Blood, 67 (1986), pp. 56–63.

[53] T. EMONET, C. M. MACAL, M. J. NORTH, C. E. WICKERSHAM, AND P. CLUZEL, *AgentCell: A digital single-cell assay for bacterial chemotaxis*, Bioinformatics, 21 (2005), pp. 2714–2721.

[54] H. N. ENGLYST AND J. H. CUMMINGS, *Digestion of the polysaccharides of some cereal foods in the human small intestine*, American Journal of Clinical Nutrition, 42 (1985), pp. 778–787.

[55] L. FAIVRE, V. PARIETTI, F. SIÑERIZ, S. CHANTEPIE, M. GILBERT-SIRIEIX, P. ALBANESE, J. LARGHERO, AND V. VANNEAUX, *In vitro and in vivo evaluation of cord blood hematopoietic stem and progenitor cells amplified with glycosaminoglycan mimetic*, Stem Cell Research & Therapy, 7 (2016), p. 3.

[56] H. F. FARIN, I. JORDENS, M. H. MOSA, O. BASAK, J. KORVING, D. V. TAURIELLO, K. DE PUNDER, S. ANGERS, P. J. PETERS, M. M. MAURICE, AND H. CLEVERS, *Visualization of a short-range Wnt gradient in the intestinal stem-cell niche*, Nature, 530 (2016), pp. 340–343.

[57] H. F. FARIN, J. H. VAN ES, AND H. CLEVERS, *Redundant sources of Wnt regulate intestinal stem cells and promote formation of paneth cells*, Gastroenterology, 143 (2012).

[58] D. FAUST, I. DOLADO, A. CUADRADO, F. OESCH, C. WEISS, A. R. NEBREDA, AND C. DIETRICH, *p38α MAPK is required for contact inhibition*, Oncogene, 24 (2005), pp. 7941–7945.

[59] J. E. FERRELL, T. Y. C. TSAI, AND Q. YANG, *Modeling the cell cycle: Why do certain circuits oscillate?*, 2011.

[60] A. G. FLETCHER, C. J. BREWARD, AND S. JONATHAN CHAPMAN, *Mathematical modeling of monoclonal conversion in the colonic crypt*, Journal of Theoretical Biology, 300 (2012), pp. 118–133.

[61] A. G. FLETCHER, P. J. MURRAY, AND P. K. MAINI, *Multiscale modelling of intestinal crypt organization and carcinogenesis*, Mathematical Models and Methods in Applied Sciences, 484 (2015), pp. 546–549.

[62] A. G. FLETCHER, M. OSTERFIELD, R. E. BAKER, AND S. Y. SHVARTSMAN, *Vertex models of epithelial morphogenesis*, Biophysical Journal, 106 (2014), pp. 2291–2304.

[63] S. A. FRANK, *History of Theories*, Dynamics of Cancer. Incidence, Inheritance, and Evolution, (2007), pp. 1–378.

[64] S. FRE, M. HUYGHE, P. MOURIKIS, S. ROBINE, D. LOUVARD, AND S. ARTAVANIS-TSAKONAS, *Notch signals control the fate of immature progenitor cells in the intestine*, Nature, 435 (2005), pp. 964–968.

[65] J. GALLE, M. LOEFFLER, AND D. DRASDO, *Modeling the effect of deregulated proliferation and apoptosis on the growth dynamics of epithelial cell populations in vitro*, Biophysical Journal, 88 (2005), pp. 62–75.

[66] A. V. GETT AND P. D. HODGKIN, *A cellular calculus for signal integration by T cells*, Nature Immunology, 1 (2000), pp. 239–244.

[67] J. A. GLAZIER, *Simulation of the differential adhesion driven |[nbsp]| rearrangement of biological cells*, Phys. Rev. |[nbsp]| E, 47 (1993), pp. 2128–2154.

[68] S. GODWIN, D. WARD, E. PEDONE, M. HOMER, A. G. FLETCHER, AND L. MARUCCI, *An extended model for culture-dependent heterogenous gene expression and proliferation dynamics in mouse embryonic stem cells*, npj Systems Biology and Applications, 3 (2017), p. 19.

[69] A. GOÑI-MORENO, M. AMOS, AND F. DE LA CRUZ, *Multicellular Computing Using Conjugation for Wiring*, PLoS ONE, 8 (2013).

[70] M. J. GONZALEZ, J. R. MIRANDA-MASSARI, M. J. BERDIEL, J. DUCONGE, J. L. RODRÍGUEZ-LÓPEZ, R. HUNNINGHAKE, AND V. J. COBAS-ROSARIO, *High dose intraveneous Vitamin C and chikungunya fever: A case report*, Journal of Orthomolecular Medicine, 29 (2014), pp. 154–156.

[71] A. GÖRGENS, S. RADTKE, M. MÖLLMANN, M. CROSS, J. DÜRIG, P. A. HORN, AND B. GIEBEL, *Revision of the Human Hematopoietic Tree: Granulocyte Subtypes Derive from Distinct Hematopoietic Lineages*, Cell Reports, 3 (2013), pp. 1539–1552.

[72] T. E. GOROCHOWSKI, A. MATYJASZKIEWICZ, T. TODD, N. OAK, K. KOWALSKA, S. REID, K. T. TSANEVA-ATANASOVA, N. J. SAVERY, C. S. GRIERSON, AND M. DI BERNARDO, *BSim: An agent-based tool for modeling bacterial populations in systems and synthetic biology*, PLoS ONE, 7 (2012).

[73] GRANER AND GLAZIER, *Simulation of biological cell sorting using a two-dimensional extended Potts model.*, Phys Rev Lett, 69 (1992), pp. 2013–2016.

[74] A. GREGORIEFF AND H. CLEVERS, *Wnt signaling in the intestinal epithelium : from endoderm to cancer Wnt signaling in the intestinal epithelium : from endoderm to cancer*, Genes & Development, 19 (2005), pp. 877–890.

[75] A. GROVER, E. MANCINI, S. MOORE, A. J. MEAD, D. ATKINSON, K. D. RASMUSSEN, D. O'CARROLL, S. E. W. JACOBSEN, AND C. NERLOV, *Erythropoietin guides multipotent hematopoietic progenitor cells toward an erythroid fate*, The Journal of Experimental Medicine, 211 (2014), pp. 181–188.

[76] P. A. HALL, P. J. COATES, B. ANSARI, AND D. HOPWOOD, *Regulation of cell number in the mammalian gastrointestinal tract: the importance of apoptosis.*, Journal of cell science, (1994).

[77] L. H. HARTWELL, J. J. HOPFIELD, S. LEIBLER, AND A. W. MURRAY, *From molecular to modular cell biology*, Nature, 402 (1999), pp. C47–C52.

[78] V. HATZIMANIKATIS, K. H. LEE, W. A. RENNER, AND J. E. BAILEY, *A mathematical model for the G1/S transition of the mammalian cell cycle*, Biotechnology Letters, 17 (1995), pp. 669–674.

[79] E. D. HAWKINS, M. HOMMEL, M. L. TURNER, F. L. BATTYE, J. F. MARKHAM, AND P. D. HODGKIN, *Measuring lymphocyte proliferation, survival and differentiation using CFSE time-series data*, Nature Protocols, 2 (2007), pp. 2057–2067.

[80] E. D. HAWKINS, M. L. TURNER, M. R. DOWLING, C. VAN GEND, AND P. D. HODGKIN, *A model of immune regulation as a consequence of randomized lymphocyte division and*

death times., Proceedings of the National Academy of Sciences of the United States of America, 104 (2007), pp. 5032–7.

[81] K. HAYASHI, S. M. CHUVA, D. S. LOPES, F. TANG, AND K. LAO, *Europe PMC Funders Group Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states*, Cell Stem Cell, 3 (2013), pp. 391–401.

[82] G. HELMLINGER, P. A. NETTI, H. C. LICHTENBELD, R. J. MELDER, AND R. K. JAIN, *Solid stress inhibits the growth of multicellular tumor spheroids.*, Nat Biotechnol, 15 (1997), pp. 778–783.

[83] M. HERBERG, I. GLAUCHE, T. ZERJATKE, M. WINZI, F. BUCHHOLZ, AND I. ROEDER, *Dissecting mechanisms of mouse embryonic stem cells heterogeneity through a model-based analysis of transcription factor dynamics*, Journal of the Royal Society Interface, 13 (2016).

[84] M. HERBERG, T. KALKAN, I. GLAUCHE, A. SMITH, AND I. ROEDER, *A Model-Based Analysis of Culture-Dependent Phenotypes of mESCs.*, PloS one, 9 (2014), p. e92496.

[85] D. J. HIGHAM., *An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations*, SIAM Review, 43 (2001), pp. 525–546.

[86] M. HOLCOMBE, S. ADRA, M. BICAK, S. CHIN, S. COAKLEY, A. I. GRAHAM, J. GREEN, C. GREENOUGH, D. JACKSON, M. KIRAN, S. MACNEIL, A. MALEKI-DIZAJI, P. MCMINN, M. POGSON, R. POOLE, E. QWARNSTROM, F. RATNIEKS, M. D. ROLFE, R. SMALLWOOD, T. SUN, AND D. WORTH, *Modelling complex biological systems using an agent-based approach*, Integrative Biology, 4 (2012), pp. 53–64.

[87] D. HOLLANDER, *The intestinal permeability barrier: A hypothesis as to its regulation and involvement in crohn's disease*, Scandinavian Journal of Gastroenterology, 27 (1992), pp. 721–726.

[88] M. HOMMEL AND P. D. HODGKIN, *TCR Affinity Promotes CD8+ T Cell Expansion by Regulating Survival*, The Journal of Immunology, 179 (2007), pp. 2250–2260.

[89] W. HONG AND K. L. GUAN, *The YAP and TAZ transcription co-activators: Key downstream effectors of the mammalian Hippo pathway*, Seminars in Cell and Developmental Biology, 23 (2012), pp. 785–793.

[90] K. HÜBNER, G. FUHRMANN, L. K. CHRISTENSON, J. KEHLER, R. REINBOLD, R. DE LA FUENTE, J. WOOD, J. F. STRAUSS, M. BOIANI, AND H. R. SCHÖLER, *Derivation of oocytes from mouse embryonic stem cells*, Science, 300 (2003), pp. 1251–1256.

[91] M. ILYAS, *Wnt signalling and the mechanistic basis of tumour development*, 2005.

[92] M. IMAJO, K. MIYATAKE, A. IIMURA, A. MIYAMOTO, AND E. NISHIDA, *A molecular mechanism that links Hippo signalling to the inhibition of Wnt / β-catenin signalling*, EMBO Journal, 31 (2012), pp. 1109–1122.

[93] T. INGHAM-DEMPSTER, B. CORFE, AND D. WALKER, *A cellular based model of the colon crypt suggests novel effects for Apc phenotype in colorectal carcinogenesis*, Journal of Computational Science, 24 (2018), pp. 125–131.

[94] T. INGHAM-DEMPSTER, D. C. WALKER, AND B. M. CORFE, *An agent-based model of anoikis in the colon crypt displays novel emergent behaviour consistent with biological observations*, Royal Society Open Science, 4 (2017), p. 160858.

[95] D. E. JAALOUK AND J. LAMMERDING, *Mechanotransduction gone awry*, 2009.

[96] S. S. JANG, K. T. OISHI, R. G. EGBERT, AND E. KLAVINS, *Specification and simulation of synthetic multicelled behaviors*, ACS Synthetic Biology, 1 (2012), pp. 365–374.

[97] R. JOHNSON AND G. HALDER, *The two faces of Hippo: Targeting the Hippo pathway for regenerative medicine and cancer treatment*, Nature Reviews Drug Discovery, 13 (2013), pp. 63–79.

[98] S. KAUFFMAN AND J. J. WILLE, *The mitotic oscillator in Physarum polycephalum*, Journal of Theoretical Biology, 55 (1975).

[99] S. K. KAY, H. A. HARRINGTON, S. SHEPHERD, K. BRENNAN, T. DALE, J. M. OSBORNE, D. J. GAVAGHAN, AND H. M. BYRNE, *The role of the Hes1 crosstalk hub in Notch-Wnt interactions of the intestinal crypt*, PLoS Computational Biology, 13 (2017).

[100] A. KEL, I. DEINEKO, O. KEL-MARGOULIS, E. WINGENDER, AND V. RATNER, *Modeling of Gene Regulatory Network of Cell Cycle Control. Role of E2F Feedback Loops*, Proceedings of the German Conference on Bioinformatics, (2000).

[101] N.-G. KIM, E. KOH, X. CHEN, AND B. M. GUMBINER, *E-cadherin mediates contact inhibition of proliferation through Hippo signaling-pathway components*, Proceedings of the National Academy of Sciences, 108 (2011), pp. 11930–11935.

[102] M. KOULNIS, E. PORPIGLIA, D. HIDALGO, AND M. SOCOLOVSKY, *Erythropoiesis: from molecular pathways to system properties. Advances in experimental medicine and biology*, Dvances in Experimental Medicine and Biology, 844 (2014), pp. 37–58.

[103] J. U. KREFT, G. BOOTH, AND J. W. WIMPENNY, *BacSim, a simulator for individual-based modelling of bacterial colony growth*, Microbiology, 144 (1998), pp. 3275–3287.

[104] H. LAMLUM, A. PAPADOPOULOU, M. ILYAS, A. ROWAN, C. GILLET, A. HANBY, I. TALBOT, W. BODMER, AND I. TOMLINSON, *APC mutations are sufficient for the growth of early colorectal adenomas*, Proceedings of the National Academy of Sciences, 97 (2000), pp. 2225–2228.

[105] A. J. LANGLANDS, A. A. ALMET, P. L. APPLETON, I. P. NEWTON, J. M. OSBORNE, AND I. S. NÄTHKE, *Paneth Cell-Rich Regions Separated by a Cluster of Lgr5+ Cells Initiate Crypt Fission in the Intestinal Stem Cell Niche*, PLoS Biology, 14 (2016).

[106] L. A. LARDON, B. V. MERKEY, S. MARTINS, A. DÖTSCH, C. PICIOREANU, J. U. KREFT, AND B. F. SMETS, *iDynoMiCS: Next-generation individual-based modelling of biofilms*, Environmental Microbiology, 13 (2011), pp. 2416–2434.

[107] E. LEE, A. SALIC, R. KRÜGER, R. HEINRICH, AND M. W. KIRSCHNER, *The roles of APC and axin derived from experimental and theoretical analysis of the Wnt pathway*, PLoS Biology, 1 (2003), pp. 116–132.

[108] H. Y. LEE AND A. S. PERELSON, *Modeling T cell proliferation and death in vitro based on labeling data: Generalizations of the Smith-Martin cell cycle model*, Bulletin of Mathematical Biology, 70 (2008), pp. 21–44.

[109] Y. LEE, S. KOUVROUKOGLOU, L. MCINTIRE, AND K. ZYGOURAKIS, *A cellular automaton model for the proliferation of migrating contact- inhibited cells*, Biophysical Journal, 69 (1995), pp. 1284–1298.

[110] P. V. LIEDEKERKE, J. NEITSCH, T. JOHANN, K. ALESSANDRI, P. NASSOY, AND D. DRASDO, *Quantitative agent-based modeling reveals mechanical stress response of growing tumor spheroids is predictable over various growth conditions and cell lines*, bioRxiv, (2018), p. 122614.

[111] N. A. LOBO, Y. SHIMONO, D. QIAN, AND M. F. CLARKE, *The Biology of Cancer Stem Cells*, 2007.

[112] B. C. LOW, C. Q. PAN, G. V. SHIVASHANKAR, A. BERSHADSKY, M. SUDOL, AND M. SHEETZ, *YAP/TAZ as mechanosensors and mechanotransducers in regulating organ size and tumor growth*, FEBS Letters, 588 (2014), pp. 2663–2670.

[113] A. L. MACLEAN, Z. ROSEN, H. M. BYRNE, AND H. A. HARRINGTON, *Parameter-free methods distinguish Wnt pathway models and guide design of experiments*, Proceedings of the National Academy of Sciences, 112 (2015), pp. 2652–2657.

[114] H. MARKS, R. MENAFRA, S. DENISSOV, AND K. JONES, *The Transcriptional and Epigenetic Foundation of Ground State Pluripotency*, Cell, 149 (2012), pp. 590–604.

[115] G. R. MARTIN, *Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells.*, Proceedings of the National Academy of Sciences, 78 (1981), pp. 7634–7638.

[116] E. MARTZ AND M. S. STEINBERG, *The role of cell‚Äêcell contact in ‚Äúcontact‚Äù inhibition of cell division: A review and new evidence*, Journal of Cellular Physiology, 79 (1972), pp. 189–210.

[117] F. D. MAST, A. V. RATUSHNY, AND J. D. AITCHISON, *Systems cell biology*, 2014.

[118] A. MATYJASZKIEWICZ, G. FIORE, F. ANNUNZIATA, C. S. GRIERSON, N. J. SAVERY, L. MARUCCI, AND M. DI BERNARDO, *BSim 2.0: An Advanced Agent-Based Cell Simulator*, ACS Synthetic Biology, 6 (2017), pp. 1969–1972.

[119] S. J. MCGEE, A. M. HAVENS, Y. SHIOZAWA, Y. JUNG, AND R. S. TAICHMAN, *Effects of erythropoietin on the bone microenvironment*, Growth Factors, 30 (2012), pp. 22–28.

[120] F. A. MEINEKE, C. S. POTTEN, AND M. LOEFFLER, *Cell migration and organization in the intestinal crypt using a lattice-free model*, Cell Proliferation, 34 (2001), pp. 253–266.

[121] H. MIAO, X. JIN, A. S. PERELSON, AND H. WU, *Evaluation of Multitype Mathematical Models for CFSE-Labeling Experiment Data*, Bulletin of Mathematical Biology, 74 (2012), pp. 300–326.

[122] G. R. MIRAMS, C. J. ARTHURS, M. O. BERNABEU, R. BORDAS, J. COOPER, A. CORRIAS, Y. DAVIT, S. J. DUNN, A. G. FLETCHER, D. G. HARVEY, M. E. MARSH, J. M. OSBORNE, P. PATHMANATHAN, J. PITT-FRANCIS, J. SOUTHERN, N. ZEMZEMI, AND D. J. GAVAGHAN, *Chaste: An Open Source C++ Library for Computational Physiology and Biology*, PLoS Computational Biology, 9 (2013).

[123] G. R. MIRAMS, H. M. BYRNE, AND J. R. KING, *A multiple timescale analysis of a mathematical model of the Wnt/??-catenin signalling pathway*, Journal of Mathematical Biology, 60 (2010), pp. 131–160.

[124] G. R. MIRAMS, A. G. FLETCHER, P. K. MAINI, AND H. M. BYRNE, *A theoretical investigation of the effect of proliferation and adhesion on monoclonal conversion in the colonic crypt*, Journal of Theoretical Biology, 312 (2012), pp. 143–156.

[125] Y. MIYAOKA AND A. MIYAJIMA, *To divide or not to divide: Revisiting liver regeneration*, 2013.

[126] T. MOROISHI, H. W. PARK, B. QIN, Q. CHEN, Z. MENG, S. W. PLOUFFE, K. TANIGUCHI, F. X. YU, M. KARIN, D. PAN, AND K. L. GUAN, *A YAP/TAZ-induced feedback mechanism regulates Hippo pathway homeostasis*, Genes and Development, 29 (2015), pp. 1271–1284.

[127] NHS, *Why Give Blood*.
\url{https://www.blood.co.uk/why-give-blood/}, 2017.

[128] A. NISHIYAMA, L. XIN, A. A. SHAROV, M. THOMAS, G. MOWRER, E. MEYERS, Y. PIAO, S. MEHTA, S. YEE, Y. NAKATAKE, C. STAGG, L. SHAROVA, L. S. CORREA-CERRO, U. BASSEY, H. HOANG, E. KIM, R. TAPNIO, Y. QIAN, D. DUDEKULA, M. ZALZMAN, M. LI, G. FALCO, H. T. YANG, S. L. LEE, M. MONTI, I. STANGHELLINI, M. N. ISLAM, R. NAGARAJA, I. GOLDBERG, W. WANG, D. L. LONGO, D. SCHLESSINGER, AND M. S. KO, *Uncovering Early Response of Gene Regulatory Networks in ESCs by Systematic Induction of Transcription Factors*, Cell Stem Cell, 5 (2009), pp. 420–433.

[129] R. E. NORDON, M. NAKAMURA, C. RAMIREZ, AND R. ODELL, *Analysis of growth kinetics by division tracking*, Immunology and Cell Biology, 77 (1999), pp. 523–529.

[130] M. J. NORTH, N. T. COLLIER, J. OZIK, E. R. TATARA, C. M. MACAL, M. BRAGEN, AND P. SYDELKO, *Complex adaptive systems modeling with Repast Simphony*, Complex Adaptive Systems Modeling, 1 (2013), p. 3.

[131] K. ORFORD, C. C. ORFORD, AND S. W. BYERS, *Exogenous expression of β-catenin regulates contact inhibition, anchorage-independent growth, anoikis, and radiation-induced cell cycle arrest*, Journal of Cell Biology, 146 (1999), pp. 855–867.

[132] S. H. ORKIN AND L. I. ZON, *Hematopoiesis: An Evolving Paradigm for Stem Cell Biology*, 2008.

[133] J. M. OSBORNE, *Multiscale model of colorectal cancer using the cellular Potts framework*, Cancer Informatics, 14 (2015), pp. 83–93.

[134] G. PAN AND J. A. THOMSON, *Nanog and transcriptional networks in embryonic stem cell pluripotency*, 2007.

[135] A. PARKER, O. J. MACLAREN, A. G. FLETCHER, D. MURARO, P. A. KREUZALER, H. M. BYRNE, P. K. MAINI, A. J. WATSON, AND C. PIN, *Cell proliferation within small intestinal crypts is the principal driving force for cell migration on villi*, FASEB Journal, 31 (2017), pp. 636–649.

[136] P. PATHMANATHAN AND R. A. GRAY, *Verification of computational models of cardiac electro-physiology*, International Journal for Numerical Methods in Biomedical Engineering, 30 (2014), pp. 525–544.

[137] S. PICCOLO, S. DUPONT, AND M. CORDENONSI, *The Biology of YAP / TAZ: Hippo Signaling and Beyond*, Physiological Reviews, 94 (2014), pp. 1287–1312.

[138] J. PITT-FRANCIS, P. PATHMANATHAN, M. O. BERNABEU, R. BORDAS, J. COOPER, A. G. FLETCHER, G. R. MIRAMS, P. MURRAY, J. M. OSBORNE, A. WALTER, S. J. CHAPMAN, A. GARNY, I. M. M. VAN LEEUWEN, P. K. MAINI, B. RODRÍGUEZ, S. L. WATERS, J. P. WHITELEY, H. M. BYRNE, AND D. J. GAVAGHAN, *Chaste: A test-driven approach to software development for biological modelling*, Computer Physics Communications, 180 (2009), pp. 2452–2471.

[139] C. S. POTTEN, M. KELLETT, D. A. REW, AND S. A. ROBERTS, *Proliferation in human gastrointesinal epithelium using bromodeoxyuridine in vivo: Data for different sites, proximity to a tumour, and polyposis coli*, Gut, 33 (1992), pp. 524–529.

[140] C. S. POTTEN, M. KELLETT, S. A. ROBERTS, D. A. REW, AND G. D. WILSON, *Measurement of Invivo Proliferation in Human Colorectal Mucosa Using Bromodeoxyuridine*, Gut, 33 (1992), pp. 71–78.

[141] P. REVY, M. SOSPEDRA, B. BARBOUR, AND A. TRAUTMANN, *Functional antigen-independent synapses formed between T cells and dendritic cells*, Nature Immunology, 2 (2001), pp. 925–931.

[142] T. REYA AND H. CLEVERS, *Wnt signalling in stem cells and cancer*, Nature, 434 (2005), pp. 843–850.

[143] T. REYA, S. J. MORRISON, M. F. CLARKE, AND I. L. WEISSMAN, *Stem cells, cancer, and cancer stem cells*, Nature, 414 (2001), pp. 105–111.

[144] T. J. RUDGE, P. J. STEINER, A. PHILLIPS, AND J. HASELOFF, *Computational modeling of synthetic microbial biofilms*, ACS Synthetic Biology, 1 (2012), pp. 345–352.

[145] M. SÁNCHEZ-CASTILLO, D. RUAU, A. C. WILKINSON, F. S. NG, R. HANNAH, E. DIAMANTI, P. LOMBARD, N. K. WILSON, AND B. GOTTGENS, *CODEX: A next-generation sequencing experiment database for the haematopoietic and embryonic stem cell communities*, Nucleic Acids Research, 43 (2015), pp. D1117–D1123.

[146] O. J. SANSOM, K. R. REED, A. J. HAYES, H. IRELAND, H. BRINKMANN, I. P. NEWTON, E. BATLLE, P. SIMON-ASSMANN, H. CLEVERS, I. S. NATHKE, A. R. CLARKE, AND D. J. WINTON, *Loss of Apc in vivo immediately perturbs Wnt signaling, differentiation, and migration*, Genes and Development, 18 (2004), pp. 1385–1390.

[147] G. SCHALLER AND M. MEYER-HERMANN, *Multicellular tumor spheroid in an off-lattice Voronoi-Delaunay cell model*, Physical Review E - Statistical, Nonlinear, and Soft Matter Physics, 71 (2005).

[148] Y. SCHMITZ, K. RATEITSCHAK, AND O. WOLKENHAUER, *Analysing the impact of nucleo-cytoplasmic shuttling of β-catenin and its antagonists APC, Axin and GSK3 on Wnt/β-catenin signalling*, Cellular Signalling, 25 (2013), pp. 2210–2221.

[149] M. C. SCHROEDER AND G. HALDER, *Regulation of the Hippo pathway by cell architecture and mechanical signals*, 2012.

[150] R. SCOGNAMIGLIO, N. CABEZAS-WALLSCHEID, M. C. THIER, S. ALTAMURA, A. REYES, Á. M. PRENDERGAST, D. BAUMGÄRTNER, L. S. CARNEVALLI, A. ATZBERGER, S. HAAS, L. VON PALESKE, T. BOROVIAK, P. WÖRSDÖRFER, M. A. ESSERS, U. KLOZ, R. N. EISENMAN, F. EDENHOFER, P. BERTONE, W. HUBER, F. VAN DER HOEVEN, A. SMITH, AND A. TRUMPP, *Myc Depletion Induces a Pluripotent Dormant State Mimicking Diapause*, Cell, 164 (2016), pp. 668–680.

[151] D. SCOVILLE AND X. C. HE, *Current View : Intestinal Stem Cells and*, Gastroenterology, 134 (2008), pp. 849–864.

[152] M. SEHL, H. ZHOU, J. S. SINSHEIMER, AND K. L. LANGE, *Extinction models for cancer stem cell therapy*, Mathematical Biosciences, 234 (2011), pp. 132–146.

[153] M. SHOSHKES-CARMEL, Y. J. WANG, K. J. WANGENSTEEN, B. TÓTH, A. KONDO, E. E. MASSASA, S. ITZKOVITZ, AND K. H. KAESTNER, *Erratum to: Subepithelial telocytes are an important source of Wnts that supports intestinal crypts (Nature, (2018), 557, 7704, (242-246), 10.1038/s41586-018-0084-4)*, Nature, 560 (2018), p. E29.

[154] E. SKLAR, *Software review: NetLogo, a multi-agent simulation environment*, Artificial Life, 13 (2007), pp. 303–311.

[155] J. A. SMITH AND L. MARTIN, *Do Cells Cycle?*, Proceedings of the National Academy of Sciences, 70 (1973), pp. 1263–1267.

[156] J. L. STAMOS AND W. I. WEIS, *The β-catenin destruction complex*, 2013.

[157] A. M. STEPHEN, A. C. HADDAD, AND S. F. PHILLIPS, *Passage of carbohydrate into the colon. Direct measurements in humans.*, Gastroenterology, 85 (1983), pp. 589–95.

[158] R. STEUER, *Effects of stochasticity in models of the cell cycle: From quantized cycle times to noise-induced oscillations*, Journal of Theoretical Biology, 228 (2004), pp. 293–301.

[159] J. P. SUNTER, D. R. APPLETON, M. S. DÉ RODRIGUEZ, N. A. WRIGHT, A. J. WATSON, AND M. S. DE RODRIGUEZ, *A comparison of cell proliferation at different sites within the large bowel of the mouse.*, Journal of anatomy, 129 (1979), pp. 833–42.

[160] M. SWAT, A. KEL, AND H. HERZEL, *Bifurcation analysis of the regulatory modules of the mammalian G1/S transition*, Bioinformatics, 20 (2004), pp. 1506–1511.

[161] M. H. SWAT, G. L. THOMAS, J. M. BELMONTE, A. SHIRINIFARD, D. HMELJAK, AND J. A. GLAZIER, *Multi-Scale Modeling of Tissues Using CompuCell3D*, Methods in Cell Biology, 110 (2012), pp. 325–366.

[162] A. SZABÓ AND R. M. H. MERKS, *Cellular Potts Modeling of Tumor Growth, Tumor Invasion, and Tumor Evolution*, Frontiers in Oncology, 3 (2013).

[163] H. TAKANO, H. EMA, K. SUDO, AND H. NAKAUCHI, *Asymmetric Division and Lineage Commitment at the Level of Hematopoietic Stem Cells*, The Journal of Experimental Medicine, 199 (2004), pp. 295–302.

[164] C. TAMM, S. P. GALITÓ, AND C. ANNERÉN, *A comparative study of protocols for mouse embryonic stem cell culturing*, PLoS ONE, 8 (2013).

[165] S. G. TANGYE, D. T. AVERY, E. K. DEENICK, AND P. D. HODGKIN, *Intrinsic Differences in the Proliferation of Naive and Memory Human B Cells as a Mechanism for Enhanced Secondary Immune Responses*, The Journal of Immunology, 170 (2003), pp. 686–694.

[166] THE MATHWORKS, *MATLAB 8.0 and Statistics Toolbox 8.1*, 2012.

[167] ——, *MATLAB 8.0 and Statistics Toolbox 8.1*, 2012.

[168] J. A. THOMSON, *Embryonic stem cell lines derived from human blastocysts*, Science, 282 (1998), pp. 1145–1147.

[169] D. TRANI, S. A. NELSON, B.-H. MOON, J. J. SWEDLOW, E. M. WILLIAMS, S. J. STRAWN, P. L. APPLETON, B. KALLAKURY, I. NÄTHKE, AND A. J. FORNACE, *High-Energy Particle-Induced Tumorigenesis Throughout the Gastrointestinal Tract*, Radiation Research, 181 (2014), pp. 162–171.

[170] S. VAISHNAVA, C. L. BEHRENDT, A. S. ISMAIL, L. ECKMANN, AND L. V. HOOPER, *Paneth cells directly sense gut commensals and maintain homeostasis at the intestinal host-microbial interface*, Proceedings of the National Academy of Sciences, 105 (2008), pp. 20858–20863.

[171] M. VAN DE WETERING, E. SANCHO, C. VERWEIJ, W. DE LAU, I. OVING, A. HURLSTONE, K. VAN DER HORN, E. BATLLE, D. COUDREUSE, A. P. HARAMIS, M. TJON-PON-FONG, P. MOERER, M. VAN DEN BORN, G. SOETE, S. PALS, M. EILERS, R. MEDEMA, AND H. CLEVERS, *The β-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells*, Cell, 111 (2002), pp. 241–250.

[172] I. M. VAN LEEUWEN, H. M. BYRNE, O. E. JENSEN, AND J. R. KING, *Elucidating the interactions between the adhesive and transcriptional functions of β-catenin in normal and cancerous cells*, Journal of Theoretical Biology, 247 (2007), pp. 77–102.

[173] I. M. M. VAN LEEUWEN, G. R. MIRAMS, A. WALTER, A. FLETCHER, P. MURRAY, J. OSBORNE, S. VARMA, S. J. YOUNG, J. COOPER, B. DOYLE, J. PITT-FRANCIS, L. MOMTAHAN, P. PATHMANATHAN, J. P. WHITELEY, S. J. CHAPMAN, D. J. GAVAGHAN, O. E. JENSEN, J. R. KING, P. K. MAINI, S. L. WATERS, AND H. M. BYRNE, *An integrative computational model for intestinal tissue renewal*, Cell Proliferation, 42 (2009), pp. 617–636.

[174] K. L. VANDUSSEN, A. J. CARULLI, T. M. KEELEY, S. R. PATEL, B. J. PUTHOFF, S. T. MAGNESS, I. T. TRAN, I. MAILLARD, C. SIEBEL, A. KOLTERUD, A. S. GROSSE, D. L. GUMUCIO, S. A. ERNST, Y.-H. TSAI, P. J. DEMPSEY, AND L. C. SAMUELSON, *Notch signaling modulates proliferation and differentiation of intestinal crypt base columnar stem cells*, Development, 139 (2012), pp. 488–497.

[175] X. VARELAS, B. W. MILLER, R. SOPKO, S. SONG, A. GREGORIEFF, F. A. FELLOUSE, R. SAKUMA, T. PAWSON, W. HUNZIKER, H. MCNEILL, J. L. WRANA, AND L. ATTISANO, *The Hippo Pathway Regulates Wnt/β-Catenin Signaling*, Developmental Cell, 18 (2010), pp. 579–591.

[176] N. VLADIMIROV, L. LØVDOK, D. LEBIEDZ, AND V. SOURJIK, *Dependence of bacterial chemotaxis on gradient shape and adaptation rate*, PLoS Computational Biology, 4 (2008).

[177] D. WARD, D. CARTER, M. HOMER, L. MARUCCI, AND A. GAMPEL, *Mathematical modeling reveals differential effects of erythropoietin on proliferation and lineage commitment of human hematopoietic progenitors in early erythroid culture*, Haematologica, 101 (2016), pp. 286–296.

[178] G. WEI, P. BOGDAN, AND R. MARCULESCU, *Efficient modeling and simulation of bacteria-based nanonetworks with BNSim*, IEEE Journal on Selected Areas in Communications, 31 (2013), pp. 868–878.

[179] S. Y. WONG, K. H. CHIAM, C. T. LIM, AND P. MATSUDAIRA, *Computational model of cell positioning: Directed and collective migration in the intestinal crypt epithelium*, Journal of the Royal Society Interface, 7 (2010), pp. S351–63.

[180] H. WU, U. KLINGMÖLLER, P. BESMER, AND H. F. LODISH, *Interaction of the erythropoietin and stem-cell-factor receptors*, Nature, 377 (1995), pp. 242–246.

[181] H. WU, X. LIU, AND R. JAENISCH, *Generation of Committed ErythroidBFU-E and CFU-E Progenitors Does Not Require Etythropoietin or the Etythropoietin Receptor*, Cell, 63 (2004), pp. 1–9.

[182] C.-C. YANG, H. K. GRAVES, I. M. MOYA, C. TAO, F. HAMARATOGLU, A. B. GLADDEN, AND G. HALDER, *Differential regulation of the Hippo pathway by adherens junctions and apical‚Äìbasal cell polarity modules*, Proceedings of the National Academy of Sciences, 112 (2015), pp. 1785–1790.

[183] Q.-L. YING, J. WRAY, J. NICHOLS, L. BATLLE-MORERA, B. DOBLE, J. WOODGETT, P. COHEN, AND A. SMITH, *The ground state of embryonic stem cell self-renewal*, Nature, 453 (2008), pp. 519–523.

[184] H. E. YOUNG AND A. C. BLACK, *Adult stem cells*, The Anatomical Record, 276A (2004), pp. 75–102.

[185] M. YOUNG, L. ORDONEZ, AND A. R. CLARKE, *What are the best routes to effectively model human colorectal cancer?*, 2013.

[186] S. YUI, L. AZZOLIN, M. MAIMETS, M. T. PEDERSEN, R. P. FORDHAM, S. L. HANSEN, H. L. LARSEN, J. GUIU, M. R. ALVES, C. F. RUNDSTEN, J. V. JOHANSEN, Y. LI, C. D. MADSEN, T. NAKAMURA, M. WATANABE, O. H. NIELSEN, P. J. SCHWEIGER, S. PICCOLO, AND K. B. JENSEN, *YAP/TAZ-Dependent Reprogramming of Colonic Epithelium Links ECM Remodeling to Tissue Regeneration*, 2018.

[187] Q. ZENG AND W. HONG, *The Emerging Role of the Hippo Pathway in Cell Contact Inhibition, Organ Size Control, and Cancer Development in Mammals*, 2008.

[188] B. ZHAO, B. ZHAO, X. WEI, X. WEI, W. LI, W. LI, R. S. UDAN, R. S. UDAN, Q. YANG, Q. YANG, J. KIM, J. KIM, J. XIE, J. XIE, T. IKENOUE, T. IKENOUE, J. YU, J. YU, L. LI, L. LI, P. ZHENG, P. ZHENG, K. YE, K. YE, A. CHINNAIYAN, A. CHINNAIYAN, G. HALDER, G. HALDER, Z.-C. LAI, Z.-C. LAI, K.-L. GUAN, AND K.-L. GUAN, *Inactivation of YAP oncoprotein by the Hippo pathway is involved in cell contact inhibition and tissue growth control*, Genes & development, 21 (2007), pp. 2747–2761.